DOCUMENT RESUME

ED 350 334 TM 019 093

AUTHOR	Plomp, Tjeerd, Ed.; And Others
TITLE	European Conference on Educational Research, University of Twente (Enschede, The Netherlands, June 22-25, 1992). Book of Summaries: Volumes 1, 2, and 3.
INSTITUTION	Twente Univ. of Technology, Enschede (Netherlands).
REPORT NO	ISBN-90-365-0534-8; ISBN-90-365-0535-6; ISBN-90-365-0536-4
PUB DATE	92
NOTE	1,358p.
PUB TYPE	Collected Works - Conference Proceedings (021) -- Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF11/PC55 Plus Postage.
DESCRIPTORS	*Conference Papers; Curriculum; Educational Administration; Educational Assessment; *Educational Research; Educational Technology; Elementary Secondary Education; *Foreign Countries; Higher Education; Instructional Effectiveness; Learning; Professional Education; Research Methodology; Teacher Education; Teaching Methods; Vocational Education

ABSTRACT

Three volumes present, in three chapters each, summaries of the papers delivered at the June 1992 European Conference on Educational Research. In 1992 the annual Dutch Educational Research Association conference became international. The 9 themes of the conference, arranged in 9 chapters, and the number of papers presented for each theme are as follows: (1) educational organization and management, 21 papers and 3 symposia; (2) curriculum, 30 papers and 6 symposia; (3) learning and instruction, 46 papers, 6 poster sessions, and 7 symposia; (4) educational technology, 46 papers, 6 poster sessions, and 2 symposia; (5) research methods and evaluation research technology, 38 papers, 3 poster sessions, and 4 symposia; (6) teaching and teacher training, 40 papers, 4 poster sessions, and 4 symposia; (7) vocational and professional education and training, 20 papers and 4 symposia; (8) higher education, 34 papers, 2 poster sessions, and 2 symposia; and (9) the social context of education, 28 papers, 2 poster sessions, and 1 symposium. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED350334

22 - 25 JUNE 1992

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. NELISSEN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BOOK OF SUMMARIES

EUROPEAN CONFERENCE ON EDUCATIONAL RESEARCH

UNIVERSITY OF TWENTE
DEPARTMENT OF EDUCATION
ENSCHEDE • THE NETHERLANDS

BEST COPY AVAILABLE

EDITORS: TJ. PLOMP, J.M. PIETERS, A. FETERIS

TM019093

European Conference on Educational Research

BOOK OF SUMMARIES

Volume 1

University of Twente
The Netherlands
June 22 - 25, 1992

COLOFON

Editors:
Tjeerd Plomp
Jules Pieters
Andries Feteris

Editorial assistance:
Harmen Abma
Jeroen Breman
Conny de Koning
Olivia Kramers
Renate Schraa

Cover:
Hanna Snijder

Print:
Duoprint

University of Twente
Department of Education
P.O. Box 217
7500 AE Enschede

ISBN 90-365-0534-8

Copyright © 1992 by Department of Education, University of Twente

CONTENTS

Preface
Purpose and scope of the conference, *Tjeerd Plomp* IX

Chapter 1
EDUCATIONAL ORGANISATION AND MANAGEMENT 1

Papers
Effectiveness-oriented educational organization and management: an introduction, *J. Scheerens* 3
The development of qualitative performance indicators for school self-evaluation and development planning in Scotland, *E.P. Spencer* 5
Cognitive and affective outcomes in school effectiveness research, *A. W.M. Knuver, & H. P. Brandsma,* 8
Dimensions of school differences in examination results allowing for intake achievements, *H. Goldstein, J. Rasbash, M. Yang, G. Woodhouse, & H. Pan* 11
Issues in outcomes assessment: The concept of 'value added' in the context of international education, *S. Johnson* 13
The impact of examinations: Qualitative and quantitative data in the context of an indicator system, *A.B. Ibrahim & C.T. Fitz-Gibbon* 16
Teacher appointment characteristics and pupil achievement, *C. Vermeulen, & R. Bosker* 19
Coordination in secondary school and its implications for student achievement, *B. Witziers* 21
School size, characteristics and outcomes in Dutch secondary education, *F.G.M. Kleintjes, & E.J.J. Kremers* 25
Economics of scale and scope in schools: An economic analysis by means of translog cost functions, *M. Vos* 28
Efficiency of Regional and Local Systems of education, *A. Petrikás* 31
Effective instruction and grouping of students, *G.J. Reezigt & M.G. Weide* 34
Influencing Educational Practice through Performance Indicators, *P.B. Tymms* 37
From school effectiveness to school improvement: Some observations on the state of play, *J. Gray* 40
School autonomy, personnel policy and educational innovation, *M. Vermeulen, H.J. Visser, Q.H.J.M. van Ojen, & P.N. Karstanje* 42
Contextual influences on school effectiveness: The role of school boards, *R. H. Hofman* 46
External support, school leadership and teacher behaviour in elementary education: A multi-level study, *K.M. Stokking* 49
Building a collaborative school culture: A case study of one woman in the principalship, *R. Wignall* 53
School principals and the entrepreneurial organization, *H. Jenkins, & D. Evans* 55

The role of the local education authority in promoting school effectiveness and evaluation, *K.A. Riley* 58
The effects of catholic secondary schooling on pupil performance, *P. Daly* 60

Symposia
Linking together school effectiveness knowledge and school: Improvement practice, *D. Hopkins, & D. Reynolds* 62
International studies on school effectiveness, *B.P.M. Creemers* 74
Conceptual and formal models of school effectiveness research, *J. Scheerens* 85

Chapter 2
CURRICULUM 97

Papers
Introduction to theme 2: Curriculum, *J. v.d. Akker, U. Hameyer, & J. Terwel* 99
Effects of cooperative learning and group composition in a secondary mathematics curriculum, *P. v.d. Eeden* 100
Mathematics textbooks and learning results in primary school grades 1 through 3, *M. v. d. Heuvel-Panhuizen & K. Gravemeijer* 103
Process-product relationships in composition: The case of revision, *U. Schuurs, & H. v.d. Bergh* 108
How to achieve the goals of environmental education? Considerations for the implementation of environmental education as a new curriculum, *J.W. Eberg* 111
Experiences in environmental education curriculum research and development in the Netherlands, *M. Pieters* 115
Man and the environment: The on-going curriculum development for pupils 15-17 in general education (1991-1994), M. *Landucci, H E. Lombardini, M. Renyier, & F.F. Rojero* 117
A review of the adequacy of current structural and logistic facilities for the implementation of environmental education in a sample of European countries, *W.D.S. Leal Filho* 121
Curriculum and European Dimension: What Can Pupils in Czechoslovakia Learn about The Netherlands, *J. Prucha* 125
Problems of Czech curriculum development: The foreign inspiration in the new Czech textbooks, *V. Martinkova* 128
Reframing the curriculum strategies in post-revolutionary Romania, *A. Dumitrescu* 131
Curricular Strategies For Promoting Lifelong Learning, *E. Ropo* 133
Equality through education: An evaluation of innovative projects concerning equality in schools, *K. Reisby* 135
Creating a curriculum to reduce student HIV-risk behaviours, *S. J. Stanley* 138
Swedish national assessment program for the compulsory school, *O. Oquist* 141
Teacher materials as a part of courseware packages: Current situation and desirable directions, *P. Keursten, & N. Nieveen* 143

Open learning as a curriculum innovation: a long road from idea to practice, *E. v.d. Berg* 146
The linking of curriculum materials development with in-service training from an implementation perspective, *M.G. Roes* 148
The use of classroom consultation in improving the implementation of a curriculum, *J. Snippe* 151
Activity-Based Learning in Elementary Schools: Implementing a Core Curriculum Principle in Elementary Science Education, *U. Hameyer, & J. v.d. Akker* 153
Exemplary teaching practices and science subjects curriculum reform, *W.A.J.M. Kuiper* 155
The experiental curriculum: Conceptualization and measurement of different levels of pupil experience, *R.F.A. Wierstra* 159
Formative research in curriculum studies, *D.F. Walker* 163
Database technology for curriculum comparison as applied to informatics education, *F. Mulder, A.E.N. Hacquebard, D. Smeets, & T. Veenstra-Strijland* 165
The development and evaluation of a psychology course for adult distance education, *W. Tomic* 168
Designing a curriculum evaluation for a vocational program in a large correctional institution, *M. J. Eash* 169
Appriciation of living environment among 10 to 12 years old school children, *L. Aho, A. Haapala, & U. Salonen* 172
The impact of research on the development of a national curriculum: A cautionary tale, *W. Harlen* 175
Coherence and continuity 5-16: English lessons for curriculum managers, *P. Weston* 179
Can there be a trans-Atlantic Curriculum Research and/or Curriculum Theory? *I. Westbury* 182

Symposia
Courseware in the curriculum: Demands on teachers, *J.M. Voogt, & J. Olson* 184
Searching for more effective curriculum development strategies, *J. v.d. Akker* 189
Some results from IEA's study 'Computers in Education', *T. Plomp* 193
Evaluating the implementation of English, mathematics and science in the English national curriculum: Procedures and early findings, *T. Russell* 206
Curriculum research and curriculum productivity: The contribution of curriculum to educational productivity, *W.Th.J.G. Hoeben* 210
Structure of educational research, *Tj. Plomp* 213

Chapter 3
LEARNING AND INSTRUCTION 221

Papers
Learning and instruction, *S. Dijkstra, & R-J. Simons* 223

Teaching energy and primary school children's alternative frameworks, *R. Trumper* 225
Two levels of sequencing in germ learning: cources and task levels, *H. Vos* 228
Problem solving in secondary physics and success in higher education, *M. Fergusson-Hessler & R. Taconis* 232
ELERIAT: A flexible experimental learning environment for research on induction and analogical transfer, *S-O. Tergan & U. Oestermeier* 236
Semantic network theory applied to mathematics education: A micro analyses, *G. Steiner* 239
The use of role playing in interpersonal skill training, *G. A. Holsbrink-Engels* 243
Effects of training in anamnestic interviewing skills, *S.M. Bögels & T.G.C. van Mourik* 245
Learning styles of learners at primary schools: Theoretical and empirical project, *Hana Krykorková* 249
Knowledge integration by means of optimization- and troubleshooting- task with computer simulation, *F. Pijnenborgh* 250
The effect of grade and logical reasoning ability on conceptions related to selected science concepts of fifth, eighth and eleventh grade students, *A. Yontar* 253
The effects of a constructivist teaching strategy on using lay-ideas on ionising radiation and the ability to make risk assessment, *H.M.C. Eijkelhof, A.E. van der Valk,& P.L. Lijnse* 256
The contribution of treatment and logical thinking ability on students' learning outcomes at high school level, *P. Askar, & O. Geban* 258
Promoting conceptual development in young children: The implications of the primary SPACE project on children's learning in science, *T. Russell, & L. Mc Guigan* 261
Map skills, a basic tool? *H.van Dijk* 263
Information processing in acquiring and using generalized knowledge, *B. Krause* 267
A study of self-efficacy for science problem solving in a cooperative instructional context, *M. de Fátima Chorão C. Sanches* 269
Models of instructional interaction and student's interpretations, *S. Järvelä* 272
Comenius' "Orbis Pictus" and some experiments in the current use of pictures in foreing language instruction, *I. Pychová* 276
Is meaningful always fruitfull? The sense or nonsense of exercises, *P.J. Meijer, & A.G. Sciarone* 281
Computer controlled reading instruction: Transfer mechanism in visual word recognition, *E. M. H. Assink, G. Kattenberg, & M. Schmitt* 285
Teaching strategies to poor readers, *M. Walraven, & P. Reitsma* 289
Computer-based reading and spelling practice, *V. H. P. van Daal* 293
Awareness of text structure: A comparison of three methods to assess awareness of structures of expository text, *A. Bergmans, M. J. M. Voeten, & C. A. J. Aarnoutse* 295
Task characteristics contributing to difficulty of functional reading tests, *W. de Geus, P. Reitsma, & K. de Glopper* 298
Quality of handwriting: Self-assessment as a tool for improvement, *M. Boekaerts, C. Nieuwenhuizen, & G. Seegers* 302

Functional writing assignment and knowledge acquisition, *G.M.T. Poorthuis, & W.A.M. Kok* 306
Active knowledge of numbers is limited and context-bound, *M. Milikowski, & J.J.Elshout* 310
A pilot study on the possibilities of realistic mathematics instruction in special education, *M. v.d. Heuvel-Panhuizen* 312
Study orientations and learning in mathematics by upper-secondary school students, *R. Yrjönsuuri* 317
Proof generation in geometry: A computer simulation of think-aloud protocols, *M. Ansellem, F. Levy, & A. Nguyen-Xuan* 320
How do pupils learn math concepts? *P. Forcheri, & M. T. Molfino* 323
Strategies in learning to use a computer application program, *H.H. Leemkuil* 325
Evaluation of a method of teaching skills: The Feuerstein's instrumental Enrichment technique in low-qualified adult's pre-training sessions, *E. Loarer, M. Huteau, J. Lautrey, & D. Chartier* 329
Conceptual growth and change through analogical reasoning and metaconceptual awareness, *L. Mason* 333
Goals and effects of learning strategies courses, *J. Kaldeway, & F. Korthagen* 337
Effects of self selected study strategies on fifth grade student's recall, reading comprehension and duration of study, *A. M. Erden, & M. Demirel* 342
Spontaneous and constrained strategy use in learning texts, *E. Czerniawska* 345
A proposition of evaluating a child's development in the process of primary education training, *M. Jakowicka* 348
Using eye movements and retrospective data to study the use of clever strategies in children's mental arithmics, *L. Verschaffel, E. De Corte, I. Gielen, & E. Struyf* 351
LOGO and problem solving, *W.van Dijk* 354
Knowledge based authoring support: A model based approach to sequencing, *A..v.d. Hulst* 357
Gender related differences in motivation and arithmic ability, *G. Seegers, & M. Boekaerts* 360
Dutch revision of Harter's self-perception profile for adolescents, *N.M..E. Neuwahl, & C. Groen-de Jong* 364
Integrating the affective domain into the instructional design process, *R.G. Main* 368
The self-regulated learner in solving mathematical problems, *M. Cardelle-Elawar* 371

Posters

Helping behaviour as an indication of cognitive modificability with children in special kindergarten, *G.M. van der Aalsvoort, & M. de Groot* 375
Proportional reasoning in chemistry: Effects of task variables and level of instruction, *A. Sanz, J.I. Pozo, & M.P. Perez-Echeverria* 378
Students' conceptions about conservation of matter, *M.A Gómez-Crespo & J.I. Pozo* 381
Science activities as a model of cooperative social behaviour for adolescents with conduct disorder, *B. R. Zinn, & D. I. Kohn* 384

Age related and individual differences in the solution of verbal analogies, *W. C.M. Resing. & N. Berger* 387
Cooperative learning and text comprehension: A simulation of real classroom practice, *J. Guillén, D. Rodr1guez, D. Saldaña, & I. Sola* 390

Symposia

Motivation and self regulation, *E. De Corte* 393
The role of domain-specific knowledge and expertise in learning and instruction: An introduction, *F. Dochy, & E. De Corte* 396
Teaching inductive thinking skills to retarded and normal students, *F.P. Büchel, G. Pelgrims, U. Scharnhorst, & P. Büchel* 398
Recent developments in instructional design, *P.R.J. Simons, & J. van Merriënboer* 411
Predisposition of learners and its influence on learning, *P.A. Kirschner, F. Dochy, P.R.J. Simons, & E. De Corte* 417
The relation between cognitive and.meteacognitive skills, *B.H.A.M. van Hout-Wolters, & J.J. Beishuizen* 432

PURPOSE AND SCOPE OF THE CONFERENCE

Tjeerd Plomp

Education is one of the important areas for which the process of unification of the European Community has manyimplications. The mutual recognition of diplomas, for example,will facilitate mobility of professional groups, such as teachers and other educational staff. Another example is that member countries will be increasingly interested in the quality of educational achievement, not only within their owncountry, but also across European countries, and within European comparison with other parts of the world (such as the United States of America and Japan). European countries are increasingly confronted with similar problems and questions, a process which will stand out even more strongly following the establishment of the Single Market in Europe by 1993. Within the framework of educational cooperation, encouragement of contacts between educational researchers may give added value to educational research in Europe. It is for this reason that the VOR, the Dutch Educational Research Association, decided to broaden the scope of its yearly educational research conference by organizing the ECER92, the European Conference on Educational Research.
ECER92 is intended to be a meeting place for educational researchers from Europe and other continents, a scientific forum where educational researchers have the possibility to discuss their research approaches and results.
As stated, ECER92 is the yearly Dutch educational research conference, which goes international in 1992. Whether another ECER will be organized is dependent on real European initiatives. Although within Europe many international and national associations of educational researchers are active in their respective domains, so far no umbrella association has been founded. If educational researchers in Europe are really interested in a network and in a regular conference as a meeting place, then this ECER 1992 provides a unique opportunity to let existing grassroot developments be integrated in an initiative for the future. This conference might be a catalyst in this process.

Framework of the conference

Being a general educational research conference, the organizing committee developed a framework for the conference which represents the broad domain in which educational researchers are active. The framework of ECER92 consists of nine themes, each of them elaborated by some coordinators.

1. Educational organization and management: issues in school effectiveness
The central question is to which extent organizational and managerial conditions of school functioning are needed in addition to curricular and instructional conditions in order to make schools effective.

2. Curriculum
This theme deals predominantly with curriculum problems that combine theoretical perspectives with evidence from empirical sources.

3. Learning and instruction
Papers are addressing general theories of learning and instruction, the description of tools for instructional design, and learning and instruction within domains.

4. Educational technology: instrumentation and implementation
The research reported covers a broad spectrum of perspectives relating to learning environments in which interactive, computer-related technology is a component.

5. Advances in research methodology and evaluation research
The majority of papers falls into the categories of the methodology of educational assessment, item response theory, and multilevel modelling.

6. Teaching and teacher training
This theme consists of the categories: nature of teaching and teaching to learn; beliefs, attitudes, and cognitions of teachers; teacher education and staff development; teacher training in international perspective.

7. Vocational and professional education and training
The following sectors are distinguished: initial vocational education, corporate training and education, and continuing (adult) education.

8. Higher education
Three levels are distinguished: the direct instructional and learning process (micro level), the organizational structure and processes (meso level), the relationships with other societal systems (macro level).

9. Societal context of education
Contributions from about 15 countries will be presented demonstrating that social and cultural factors, varying over these countries, will appear to have their own substantial effects.

Each of the themes is elaborated in a chapter, which also contains summaries of most of the papers presented.

The program
We consider it an honour that at the opening session the conference was addressed by dr. ir. J.M.M. Ritzen, Minister of Education and Sciences of the

Netherlands, and by Mr. D. Lenarduzzi from the Task Force on Human Resources, Education, Training and Youth of the European Community. Their contribution to the conference demonstrates that policy makers are valuing the idea of this conference as a meeting place for European educational researchers. Prof. T. Neville Postlethwaite (University of Hamburg) presented a key note address 'Cooperative educational research in Europe', drawing from his experience of many years in conducting international comparative research in the context of IEA, the International Association for the Evaluation of Educational Achievement, and the European Academia of Education.
The remainder of the conference consisted of five types of sessions: key note addresses, paper presentations, symposia and poster sessions, while two workshops will be conducted during the conference. Next to these sessions, an exhibit was organized where publishers and universities and educational research institutions could present their work.

Publication of papers
This book only contains summaries of papers presented. The conference organizers have stimulated the theme-coordinators to publish selections of papers as books of as special issues of scientific journals.

Acknowledgements
Given the importance of 1992 for the unification of the European Community, the VOR, the Dutch Association for Educational Research, decided to broaden the perspective of their yearly educational research conference. This initiative resulted in ECER92, the European Conference on Educational Research, organized by the University of Twente, coordinated from the Department of Education.
Many persons inside and outside the university were involved in the organization of this conference. Some to a larger extent than others. As chair of the organizing committee I would like to acknowledge the great efforts of the 'some' who were heavily involved in preparing ECER92:

- Andries Feteris, who coordinated all preparatory activities;
- Janny Spierenburg and Janita Haan of the University Convention Bureau BASICS;
- the theme coordinators who were responsible for the content of the program;
- the student organization TObias and the many students of the Department of Education who assisted during the conference days and did preparatory work on the program, the book of summaries, the social evening, the technical equipment, etc. Special thanks goes to Lotte de Goeij, Olivia Kramers, Jeroen Breman, Marlies van den Ende, Renate Jansen, Greg Ball, Renate Schraa, Luuk Kornelius, Gerlinde Podt, Diederik Stoel, Harmen Abma, Peter Faulhaber and Conny de Koning;
- Hanna Snijder-Gospodarek who designed the ECER-logo and who, together with Jan Nelissen, took care of the exhibitions;
- Thyra Kuiper, who assisted in the secretarial work.

We also would like to express our gratitude to all who subsidized the conference: the Dutch Ministry of Education and Sciences; the Task Force on Human Resources, Education, Training and Youth of the European Community; the Dutch

Institute of Educational Research; the City of Enschede; the University of Twente; the Department of Education; Shell; and Akzo.

Thanks to all of them ECER92 could become a success: more than 500 presentations for more than 700 participants!

Tjeerd Plomp
Chair Organizing Committee

CHAPTER 1

EDUCATIONAL MANAGEMENT AND ORGANIZATION

EFFECTIVENESS–ORIENTED EDUCATIONAL ORGANIZATION AND MANAGEMENT: AN INTRODUCTION

J. Scheerens, Department of Education, University of Twente, Enschede, The Netherlands

The collection of papers and symposia that will be presented in theme 1 reflects the explicit aim of the theme–coordinators to relate educational organization and management to effectiveness issues. In about 50 of the papers organization and management variables are related to educational outcomes. Roughly 25% deals with issues of school evaluation as an instrument for school management and the remaining 25% address issues of school improvement and educational innovation inspired by the results of empirical school effectiveness research. The majority of contributions has the school level as its focus, whereas only a few papers look into the influence on school functioning of educational policy measures and organizational arrangements at higher administrative levels.

When one compares the contributions to theme 1 of ECER with the collection of papers in division A of the 1992 convention of the American Educational Research Association there are some striking differences.

A categorization – as rough as the one just made of the ECER–contributions in theme 1 – reveals that about 40% of the AERA–division A papers dealt with "restructuring" or "reframing" of education at large and school management in particular; 35% focused on questions about school leadership; 10% was somewhat related to effectiveness issues and 15% discussed other topics.

Although there were some very interesting symposia at the 1992 AERA–conference about the contribution of economic theory (public choice theory, educational production functions) to educational administration issues the paradigm shift Boyd and Crowson foresaw in 1985 seems not yet to have taken place in American research in educational administration. The main aspects of this paradigm shift being: an outcome vs. a process perspective and the use of axiomatic (economic) theory as opposed to a more inductive and descriptive approach.

It would be altogether misleading to see the ECER-collection of papers as proof that European research in Educational Organization and Management is generally more outcome-related than American research under the heading of educational administration. Again it should be emphasized that effectiveness-related contributions where explicitly asked for in the call for participation of the ECER–conference. On the other hand the prosperity of the International Congress for School Effectiveness and School Improvement (ICSEI) in Europe may be

indicative of a change in perspective between the United States (where the label school effectiveness has definitely gone out of fashion) and Europe.
More or less from these trends within different educational research communities there are some important substantive issues related to the idea of a more outcome–oriented approach in educational organization – and management – research. Two of these are:

- The long causal chain between policy–measures and organizational arrangements on the one hand and student achievement on the other. Doubts in this respect are provocatively phrased by March in his statement that "changing education by changing educational administration is like changing the course of the Mississippi by spitting in the Allegheny".
- The fact that organizational theorists have pointed out that the productivity concept that is common in educational effectiveness research is to be seen as just one (perhaps narrow) perspective on organizational effectiveness.

It is to be expected that these issues will be addressed in the discussions the various contributions to theme 1 of the ECER will hopefully evoke. Finally, there are three orientations within the set of contributions that are worth being pointed at: First, there is an explicit international orientation present, particularly in the symposium organized by Creemers. There are at present some exciting efforts to design and conduct international comparative school effectiveness research. Secondly, there are some attempts in enstrengthening the conceptual and theoretical basis of school effectiveness research while using formal modeling approaches as a bridge between conceptual models and empirical research. This is the focus of the symposium organized by Scheerens.
Third, following the tradition of the ICSEI–conferences, attempts are being made to link the world of school effectiveness research more closely to school improvement. The symposium organized by Reynolds is directed to this aim.
Despite alternative approaches and substantive issues for debate an effectiveness–oriented approach to educational organization and management can be seen as a fruitful area of educational research.

Reference

Boyd, W.L. & R.L. Crowson (1985). The changing conception and practice of public school administration. Review of Research in Education, 9, 311–373.

THE DEVELOPMENT OF QUALITATIVE PERFORMANCE INDICATORS FOR SCHOOL SELF-EVALUATION AND DEVELOPMENT PLANNING IN SCOTLAND

E.P. Spencer, Scottish Office Education Department, Research and Intelligence Unit - Management of Educational Resources Unit, United Kingdom

Context

Her Majesty's Inspectors of Schools in Scotland (HMI) are actively engaged in the promotion of quality assurance systems for Scottish schools at 3 levels - national, regional and individual schools. HMI have long experience of in-depth evaluation of schools in their own annual inspection program of a sample of schools. As part of the recent development in Scotland of techniques for evaluating school effectiveness, they have carefully re-analyzed the nature of quality in a school to create a more coherent basis for their own inspections and to offer to schools and regional education authorities a framework for the 2 new levels of evaluation in the national educational system. This re-analysis has constituted a kind of action research in areas of the educational field in which hitherto there had been no systematic published work in Scotland and little elsewhere: the conceptual clarification of the nature of a "quality" educational experience and the creation for evaluators, whether inspectors, regional educational authority staff or teachers, of practical, usable means of judging quality and planning to improve it. This paper describes the issues addressed and the outcomes of a 3-year period of critique, analysis, development, trial in HMI inspections, revision and final preparation of a practical package of performance indicators for use in school self-evaluation.

Key Issues for Rationale

In establishing a rationale for qualitative performance indicators, 2 major issues have been explored:

- What is the nature of good learning and good teaching; and what is the most "powerful" and the most practical way of aggregating their many components in "chunks" which remain valid reflections of "what really matters" and are also manageable by HMI and teachers in focusing observations and evaluations?
- What criteria (if any) can be firmly established, and what system for measuring quality in learning and teaching and for using the evaluation positively?

Aggregation of components of good learning and teaching

The main problem was in practice the basis for aggregation of the various components. There are, of course, very significant underlying value questions if one seeks to carry out a fundamental analysis of good learning and teaching. HMI were not, however, engaged in such fundamental analysis. They were, rather, operating within a context of ideas about curriculum continuity, progression and balance, and about learning and teaching, which are sustained by a broad national consensus.

These generally accepted ideas include, for example, 3 types of "good learning" which it is intended pupils should achieve within a national curricular framework - (1) a set of general attributes which contribute to development as a competent person, ultimately with a belief in one's own values and confidence in one's own knowledge and abilities; (2) skills for acquiring, creating and using ideas and knowledge, for communicating opinions and feelings, and for exploiting the potential of media and techniques in art, music, drama and physical education; and (3) knowledge and understanding of concepts central to the various aspects of the curriculum, involving a personal grasp of what has been learned (not just ability to repeat memorized material).

In respect of "good teaching", the consensus recognizes that it should be, first, systematic, though the "system" is complex, including, for example, analysis of the structure or content of what is to be learned and also the many factors influencing a pupil's ability to learn. Good teaching is not only, however, a kind of applied science based on guiding principles drawn from research, thought about education and the activities of recognizably good teachers. It is also open-ended to some extent, allowing teachers and pupils opportunities for creativity, as well as coherent, purposeful, activities in pursuit of planned aims.

HMI experience and educationists' thinking about learning and teaching led easily to a long list of factors to be taken into account in evaluating them. These factors were too detailed and fragmented to be dealt with separately. Existing groupings of the factors - for example, the categories into which they were grouped for comment in HMI reports on schools - presented several problems. Confusion occurred in commenting on the planned curriculum and the implemented curriculum; the nature of "differentiation" of curriculum, learning and teaching to match pupils' needs was uncertain; the roles and influences of school managers and classroom teachers were not clearly distinguished; the relative significance for "quality" of provision of accommodation and resources and exploitation of these was ill-defined; so were the very concepts of learning (emphasis on the pupil's activities) and teaching (emphasis on the teacher's); the nature of effective assessment and its relationship to learning were in the process of clarification; views were changing about the relative significance of "process" or educational experience and "product" or attainment/knowledge/skills; several of these issues, and some others, tended to arise more than once in the process of inspection and reporting, in a way which suggested unnecessary duplication.

The outcome of attempts to overcome these problems has been a "framework for audit" for each of the main elements in a school (e.g., in a secondary school, subject departments, learning support provision for pupils with special educational needs, guidance and pastoral care, school management and ethos) in which a manageable number of key areas of quality have been identified and grouped, each containing a small number of related themes. The 'groupings' for a

BEST COPY AVAILABLE

secondary school department, for example, comprise Quality of Courses, Quality of Learning and Teaching, Pupil Progress and Attainment, Management (of Resources, Staff and Planning) and Ethos. Each key area under a main heading - e.g., under "Learning and Teaching", the key area Quality of the Teaching Process - is an aspect of the work of the department so important that evaluation should cover it. It is also large enough to avoid fragmentation, but small enough to be observed, considered and evaluated comfortably. Each is a qualitative performance indicator. The framework is thus manageable. It is also "powerful" in the sense that it covers, without duplication a great many factors likely to come into play in evaluating the work of a department.

Criteria and Measurement of Quality

Because teaching and learning are activities which are both systematic and creative, it is difficult, and perhaps undesirable, to identify specific criteria for quality. HMI experience over many years - and particularly in the process of developing the performance indicators - has suggested that it is best to seek evidence for the implementation of principles, rather than set up very specific "sine qua non" criteria. There may be several successful ways in which a particular principle can be put into practice. Illustrations of good practice should therefore have the status only of examples of the kind of evidence which might show that the key things are there.

A central concern was how best to measure quality in relation to each of the qualitative indicators described above. Staff development experience, and piloting of the system of performance indicators, led to the view that it is helpful to provide exemplification not only of very good performance but also of at least one less satisfactory level. HMI's original concept of 4 descriptors of exemplar performance for each indicator was abandoned as too detailed, but a "ladder" of 4 steps has been retained, with descriptions of typical practice at the second and the top rungs, as follows:

- Major strengths - a very good performance.(This level is exemplified).
- Strengths outweigh any weaknesses - some improvement desirable.
- Strengths outweighed by weaknesses - significant improvement needed.(This level is exemplified).
- Major weaknesses - unsatisfactory performance.

In using the performance indicators in school self-evaluation, teachers and senior staff are invited to compare their own practice with the exemplars and consider how effectively they are implementing the key principles for the aspect of their work they are evaluating, how they might move to a higher level in the "ladder" and what development plans they need to make to do so.

COGNITIVE AND AFFECTIVE OUTCOMES IN SCHOOL EFFECTIVENESS RESEARCH

Anja W.M. Knuver, & Hennie P. Brandsma, Centre for Applied Research on Education (OCTO), University of Twente, Enschede, The Netherlands

Introduction and research question

Traditionally school effectiveness research focuses on cognitive output measures, mainly on basic skills like language and arithmetic achievement. The aim of education however reaches beyond basic skills. Other skills and affective outcomes are found to be important too.

Few studies in school effectiveness research not only look at cognitive pupil outcomes but also at affective characteristics. Besides, the results of these studies are not always in accordance with each other. Rutter et al. (1979) report that cognitively effective schools are also effective for affective outcomes (i.c. the percentage of delinquent pupils at school). Brookover et al. (1979) measured the self concept of pupils and found a negative relationship between mean self concept and mean cognitive achievement of schools. Mortimore et al. (1988) again find a different result; the researchers find no relationship between affective outcomes (like behavior in school, attitudes towards school activities and self views) and cognitive outcomes. They conclude the two domains to be independent.

When correlations between affective and cognitive outcomes are calculated at the pupil level (and not after aggregation to school level) the results are more in agreement with each other. Mostly moderate positive effects are found (Hansford and Hattie, 1982). The causality direction of the relationship between affective and cognitive functioning of pupils however is not clear. In recent research Helmke (1989) found achievement to cause attitudes rather than the other way around, whereas in earlier studies affective pupil characteristics (next to other background variables) were assumed to cause achievement.

In this paper we address ourselves to the question of size and direction of the relationship between affective and cognitive outcomes on the pupil level. At the school level we look if schools that are effective in cognitive outcomes are also effective in affective outcomes.

Data and method
To answer the research questions data of a sample of about 7000 pupils of grade 6 and 8 (age groups 10 and 12) in 212 Dutch primary schools were used. Information is available on the following variables.

Cognitive outcomes measured by standardized tests:
- achievement in language
- achievement in arithmetic

Affective outcomes measured by questionnaires:
- attitude toward language
- attitude toward arithmetic
- achievement motivation
- academic self concept
- school well being

Back ground characteristics:
- intelligence (ISI)
- socio economic status of the parents
- gender
- nationality

At the pupil level pearson correlations are calculated between cognitive, affective and background pupil variables. A LISREL analysis was performed to make clear in which way affective and cognitive characteristics influence each other and the role background characteristics play in this. For the LISREL analysis the variables mentioned before, measured in grade 6 and 8, are used, but also a measure of achievement in and attitudes toward language and arithmetic, and achievement motivation, measured one year earlier (as a pretest).
At the school level rank numbers on each cognitive and affective variable for each school are computed by means of the VARCL computer programme. A three level analysis was conducted and the posterior means at the school level were saved for this goal. The ranknumbers represent the mean functioning of schools on the variables, controlled for the background variables, intelligence, SES, gender and nationality. Pearson correlations between school ranknumbers for cognitive and affective variables are computed.

Results
At the pupil level the results show that cognitive and affective pupil functioning are moderately positive related to each other. The developed LISREL model shows that there is a reason to believe that achievement causes affective functioning rather than the other way round.
At the school level small but positive correlations are found between school rank numbers on affective and cognitive outcomes.

Discussion
The developed LISREL model shows that there is a reason to believe that achievement causes affective functioning rather than the other way around. In most models and learning theories affective factors are considered as a

prerequisite for learning. The LISREL analyses are not in line with these ideas, but show that their cognitive achievement is on the contrary important for the way they think about learning and school, which in turn can influence the way they start future learning tasks.

At the school level we see the suggestion that schools can be effective in both the cognitive and the affective domain not denied. The correlations between the effectiveness and the 'affectiveness' are indeed small, but they are never negative. This means that we can conclude that effectiveness in the cognitive and affective domain can go together. In any case we cannot say that effectiveness in the one domain bothers the effectiveness in in the other domain. We think we can draw a more positive conclusion: effectiveness in the cognitive domain may even fortify effectiveness in the affective domain.

Literature

Brookover, W., Beady, C., Flood, P., Schweitzer, J., & Wisenbaker, J. (1979). School Social Systems and Student Achievement. Schools Can Make a Difference. New York, Bergin Publishers Book.

Hansford, B.C., & Hattie, J.A. (1982). The Relationship Between Self and Achievement/Performance Measures. Review of Educational Research, 52, 1, 123–142.

Helmke, A. (1989). Affective Student Characteristics and Cognitive Development: Problems, Pitfalls, Perspectives. International Journal of Educational Research, 13, 8.

Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1988). School Matters: The Junior Years. Somerset: Open Books.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). Fifteen Thousand Hours. Secondary Schools and Their Effects on Children. London: Open Books.

DIMENSIONS OF SCHOOL DIFFERENCES IN EXAMINATION RESULTS ALLOWING FOR INTAKE ACHIEVEMENTS

H. Goldstein, J. Rasbash, M. Yang, G. Woodhouse, & H. Pan, United Kingdom

Recent developments in multilevel modelling allow us to fit models where the response is a proportion, using a logit transformation. This is especially useful in many kinds of study of school effectiveness, for example where the outcome of interest is failure or success in an examination or in selection for employment or higher education.

Consider the simple model with an explanatory variable (X_1) at the student level 1 and an explanatory variable (X_2) measured at the school level 2 and a response which is binary, that is zero if a student fails the exam and one if she or he passes. We first write the model for the probability of success (π_{ij}) for the i - *th* student in the j - *th* school as follows

$$\log\left\{\frac{\pi_{ij}}{1-\pi_{ij}}\right\} = \beta_0 + \beta_1 x_{1ij} + \beta_2 x2_j + u_j \qquad (1)$$

Here u_j is a simple school level residual or 'effect' and we will generally be interested in the variance of this term and in making 'posterior' estimates, j, for each school, together with standard errors. The observed responses consist of a O or 1 for each student and we would normally assume that these have a binomial distribution with mean value the probability given by (1), although it is straightforward to fit a model in which 'extra-binomial' variation occurs. A detailed discussion of the properties of this model and how to obtain suitable parameter estimates is given by Goldstein (1990).

The paper fits an extended model of this kind to a data set consisting of exam success in public examinations at 16 years with measurements on over 10,000 students from 130 secondary schools in London. Data are available on the intake achievements of the students, their ethnic background, and gender. In addition, there is school level information on the average 'poverty' level of the school, the proportion of children in ethnic minority groups, and the religious and gender composition of the school.

The analysis can be elaborated beyond the variance components model (1) to study the incorporation of random coefficients. In particular, the paper will study

whether differences between ethnic minority groups vary across schools, and whether school differences are related to the intake achievements of the students. The paper will study estimates of school 'effects' and discuss the extent to which these can convey useful information about individual schools. The paper will also discuss the relevance of this kind of analysis for educational policy and the problems of using the results to inform people in general. Attention will be focussed particularly on the need to make adequate adjustments for intake achievement and the need to take account of the standard errors of the posterior estimates.

The data analysis is carried out using the software system ML3 (Prosser et al, 1991), using a set of specially written macros for handling general models of this kind and Paterson (1991) gives details of how to set up such a model within ML3.

References

Goldstein, H. (1991). Nonlinear multilevel models, with an application to discrete response data. Biometrika, 78, 45-51

Prosser, R., Rasbash, J. and Goldstein, H. (1991). ML3: Software for three-level analysis. London: Institute of Education.

Paterson, L. (1991). Multilevel Logistic Regression. In R. Prosser, J. Rasbash, and H. Goldstein (eds.). Data Analysis with ML3. London: Institute of Education.

ISSUES IN OUTCOME ASSESSMENT: THE CONCEPT OF 'VALUE ADDED' IN THE CONTEXT OF INTERNATIONAL EDUCATION

Sandra Johnson, Franklin College, Switzerland

Until very recently efforts to evaluate institutional effectiveness within higher education in the US have focused on resource enquiries and on student questionnaire surveys of various kinds. Physical plant resources, library resources, computer resources, financial resources, faculty credentials, all these are noted and judged as satisfactory or otherwise for the purpose of education in the fields offered and to the levels required. In addition, students are regularly invited to respond to questionnaires exploring their opinions about the service provided them by their institutions, with regard both to academic and to social provision. When they eventually leave their institutions, they continue to be invited to collaborate in institution-organized surveys, aimed to gather information about their progress in subsequent graduate education and careers, and to explore their views about the value to them of their earlier college experience.

The effectiveness of institutions in terms of their influence on students' cognitive development has not generally received the same degree of attention. Like elsewhere, there have always been commonly accepted rankings among institutions within US higher education, at least partially based on the quality of student input as indicated by average SAT scores. In other words, relative institutional quality with regard to the transmission of knowledge and the development of cognitive abilities and skills, as in other countries around the world, has been assumed indicated by input measures rather than by output measures.

Recently, though, in sympathy with renewed widespread concern about educational standards at every level of education in the US, initiatives have been taken which will change the situation. Institutions within higher education are being asked to embrace post-entry cognitive measures alongside the traditional self-reported affective measures in institutional evaluation. Even more interestingly, at least one of the six regional accrediting bodies is encouraging, if not demanding, its institutions to attempt to measure changes in both areas during the course of a student's education within the institution, and to attempt to attribute some or all of this change to the influence of that education (Middle States Commission on Higher Education 1990). From the combined research efforts directed at school effectiveness in Europe and elsewhere over very many years, it can be anticipated that attempts to measure 'value added' will prove problematic to say the least.

For many US-accredited institutions functioning outside of the US, the challenge to produce interpretable statements is a particularly interesting one to contemplate. This is more true now than might have been the case even five years ago, given some radical changes in the composition of student bodies. Some of these institutions have traditionally recruited a high proportion of US students, either as temporary 'study abroad' juniors or as 2-year program students intending to transfer stateside to embark on their majors, or as 4-year BA students. Over recent years, various factors have served to reduce steadily the number of US students willing to study abroad: principally demographic trends and economic recession in the US itself, and political instability around the globe. The reduction in US student numbers has been more than balanced in many cases by an increase in student numbers from everywhere else in the world, and in particular from Europe. The introduction of degrees in business studies has proved a specially important factor in this compositional change.

Thus, institutions which just five years ago would have recorded a proportion of US students at 75% or higher are now reporting proportions as low as 25% or less. This increasing 'internationalization' of US colleges abroad would not in itself introduce additional complications for outcomes assessment if it were the case that the students had nevertheless received a US-style secondary education and had attempted the SAT or equivalent as all college applicants stateside are expected to do. However, this is generally not the case.

The majority of the non-US students recruited into US colleges abroad are being drawn directly from the secondary school systems of their own countries, and they are entering college on the strength of their own school leaving certificates (for which they usually receive up to one year of credit). They are not required to take the SAT, and most do not. These institutions are consequently unable to appeal to average SAT scores for a handy relative reference in terms of initial student quality, as stateside institutions do (cf. Fincher 1990). Institutions recruiting from around the world are therefore faced with tremendous problems of certificate comparability if they are to use school leaving certificates as an indication of student quality on entry (Johnson 1990); existing guidelines on comparability for advanced standing decisions (such as Sepmeyer 1986) being too crudely-based for this purpose.

Whether institutions choose to face this particular comparability problem or not, they will have no choice but to accept the challenge of developing entry tests reflecting their particular mission statements - and tests which are fair to all student entrants irrespective of their various native languages. This is rendered particularly necessary by the nature of the curriculum, which is, of course, credit based and offers students an unusual flexibility in terms of course choice in the first two years. The typical variety of courses taken by students, combined often with sometimes very small class sizes, means that it is difficult, impossible even, to use course grades in any very meaningful way with respect to general institutional goals, although they obviously have a role to play in evaluating course-based goals.

For some colleges abroad, an added complication is to be found in mission statements. These often embrace goals which have rarely, if ever, been researched in any depth before: 'to promote increased tolerance and understanding of other cultures', for example. Like all other institutions engaged in

outcomes assessment, the first task for those operating in the sphere of international education is to translate mission statements into variables which might in principle be measurable, and to agree change criteria which might be practically applicable. This alone presents rather interesting challenges, inviting thought, analysis and debate about the very meaning of the term 'international education'. And this before all the more familiar difficulties documented by those with prior experience in institutional effectiveness research, and touched on above, are even contemplated.

Despite the inevitable difficulties which will be faced, indeed one might say because of these, the exercise will undoubtedly move forward the frontiers of knowledge in the field of educational evaluation. The presentation focuses on the general issue of outcomes assessment in the context of international education, giving special attention to the notion of 'value added', and using a case study for illustration of particular points of interest.

References

Fincher, C., (1990). Social justice, public interest and the SAT, in: Luijten, A.J.M. (ed.), Issues in public examinations, Utrecht: Lemma.

Johnson, S., (1990). Determining equivalence among European educational qualifications, in: Luijten, A.J.M. (ed.), Issues in public examinations, Utrecht: Lemma.

Middle States Commission on Higher Education, (1990). Designs for Excellence.

Sepmeyer, I.H. (ed.), 1986, The Country Index, Revised Edition, Vol.1, North Hollywood, CA, International Education Research Foundation, Inc.

THE IMPACT OF EXAMINATIONS:QUALITATIVE AND QUANTITATIVE DATA IN THE CONTEXT OF AN INDICATOR SYSTEM

A.B. Ibrahim & C.T. Fitz-Gibbon, The University of Newcastle Upon Tyne, United Kingdom

Introduction

Externally set and marked examinations are a long-standing feature of secondary education in the United Kingdom. They have nevertheless come under attack on numerous occasions and various educators have called for them to be replaced with other forms of assessment (Murphy and Torrance, Hargreaves, Nuttall, Burstall). Meanwhile there is a growing demand in the US for alternatives to the computer-marked multiple choice tests which have been almost the only measures available of the cognitive outcomes of schooling. The alternatives are referred to variously as 'authentic', 'complex', and/or 'performance-based' assessments (Resnick, 1990; Linn, Baker and Dunbar, 1991).

The Advanced Level examinations ('A-levels'), taken by UK students at about 18 years of age, can arguably be regarded as authentic, complex, performance-based assessments. A study of A-levels has therefore more than a parochial UK interest.

A-levels have been the target of many reform proposals. They have so far survived and are often referred to as 'the gold standard'. A single study cannot resolve issues of long term policy but it can contribute to keeping rhetoric in check by providing grounded and triangulated accounts of actual practice.

Data resouces

One source of data was the A-Level Information System (ALIS). In addition to *cognitive variables*(ability, prior achievement, examination results) this system provided extensive information about students' *attitudes* based on a questionnaires administered by university personnel to approximately 7,000 A-level candidates in over 120 schools and colleges. The questionnaire also contained students' estimates of the frequency of use of *various teaching-learning activities* (e.g. dictated notes; student presentations to the class.) These estimates represented measures of *process variables*, in particular 'alterable' process variables.

This extensive database from the ALIS project was supplemented by other measures obtained in the course of an in-depth study of two A-level subjects (Mathematics and Biology) in six schools. Twelve classrooms were studied over period of two years using observations, interviews and student diaries.

Research questions

In this paper the following questions will be addressed:

- To what extent did the forthcoming examination exert an influence over the teaching of A-level subjects? Was there evidence, for example, that the examination was leading teachers to adopt a highly didactic approach? When didactic approaches were adopted were they effective?
- To what extent did the forthcoming examination appear to affect students? For example, were there signs of great anxiety? Were students helpful towards each other as they faced the competitive examination? Were they willing to seek help from teachers?
- What influences did some structural variables appear to exert on classroom practice? For example, was class size a constraint on certain teaching-learning activities? Did the ability-composition of the class have discernible effects in either the qualitative or quantitative data?
- What hypotheses regarding effective methods of teaching Mathematics and Biology can be derived from this data set?

Results

Research question 1

Throughout the two years of preparation for the A-level examinations, every teacher was acutely aware of the exams and directed the teaching almost exclusively to the content in the examination syllabus. However, this practice appeared to be much more a matter of 'curriculum alignment' rather than 'teaching to the test'. Furthermore, within the constraint of covering the syllabus, different teachers adopted quite different styles; there was no evidence that only a didactic approach was adopted because of the perceived importance and pressure of the examination. Nor was there evidence that frequent use of examination-dominated activities (such as item-spotting, giving out extensive duplicated notes or employing more than the usual amount of examination practice) was associated with better residuals (i.e. with greater student progress or with better than predictable examination results).

Research question 2

As with teachers, students were very aware of the forthcoming examination. Anxieties were expressed but did not seem to be at levels beyond the coping strategies of students. (Those for whom the examinations were an intolerable prospect may of course already have opted out of the system.) Among the coping strategies were help-seeking behaviours. These were observed in classrooms and reported in interviews. The examination did not deter help-seeking and help-giving among students but, as might be predicted from social-psychological theories, 'evaluation apprehension' did affect help-seeking from teachers. Offered help was not so threatening to self-esteem as requested help but for several students who were falling behind, help was rarely offered and rarely requested.

Patterns of help-giving from teachers were very different between Mathematics and Biology. In the former a most every lesson provided chances for individual attention whereas such attention was generally only available in Biology in out-of-class hours.

Different teachers appeared to adopt different policies in offering help, some being almost entirely reactive to student demand and others more proactive. Their styles of helping differed also, some continuing a didactic rather than an interactive approach even in a one-to-one situation.

Research question 3

In observation and interview data, and in students' diaries, there was evidence that both class size and the perceived abilities of the students affected teacher behaviour. In particular teachers felt deterred from using certain techniques in the situation of a large class (eg student presentations) and teachers felt less able to give help to students who were not making sufficient efforts or who were perceived as unlikely to be successful in the examination for other reasons. However, there appears to be little evidence in the large body of quantitative data that, in general, residuals varied substantially with class size. (Analyses are continuing on the influence of class size on the teaching techniques employed. They will be completed before the conference.)

Research question 4

One myth for which yet another data set shows no evidence is the myth of teacher expectations. On the contrary, the more pessimistic teacher obtained the most positive residuals.
(Further analyses are in progress and will draw not only on 1991 data but also on effectiveness data for the same teachers in the previous two years)

Discussion

There is room for improvement in any system but it will be argued that the externally set and graded, complex, 'examination board', system of examining has much to recommend it. When incorporated into an information system which permits evaluation of the results in context, as progress rather than simply as raw outcomes, the examinations form a bed-rock for a quality assurance system. As Walker and Schaffarzick concluded in the 1960s from extensive studies of the then new science curricula, the important influence on achievement is 'content inclusion and emphasis'. The examination board system ensures that teachers are constantly guided vis-a-vis content inclusion and emphasis. The course which a student receives is therefore basically the same regardless of the school attended. This careful coverage of the syllabus represents one aspect of teacher accountability that Resnick et al. are seeking in setting up Examination systems in the US. It appeared to be an accountability and responsibility thoroughly and professionally accepted by teachers. Indeed, it is primarily school staff who have promoted the spread of the monitoring system, ALIS.
That these important A-level examinations, the entry-requirements for higher education and the professions, showed correlations of almost negligible size (0.10) with home background measures, was a finding which speaks in favour of the examination system and the benefits it can bring to disadvantaged areas.

TEACHER APPOINTMENT CHARACTERISTICS AND PUPIL ACHIEVEMENT

Cees Vermeulen, & Roel Bosker, Centre for Applied Research on Education (OCTO), University of Twente, Enschede, The Netherlands

Introduction

In 1985 a new law was introduced in the Netherlands, governing primary education. This law entailed the merging of kindergarten schools and primary schools in order to form one continuous form of education: the new 'basisschool'. Another aspect of this law was the fact that all teachers in the new primary education were, from that moment on, legally qualified to teach children of all age groups, regardless of their professional training. Since teachers who work with the 4-5 year old received, up to that date, a different training then teachers who work with the 6 to 12 year old, this meant that teachers could be called upon to teach children for the teaching of which their professional training did not equip them. This phenomenon is referred to as the' complete usability of teachers'. The situation in which school principals are forced to use teachers in this way is mostly brought about by a change in the number of pupils, as well as by the need to replace teachers who are sick or who are of duty by reason of short time working. The new law has thus increased the options a school principal has to deploy teachers throughout the school. Combined with the effect of an growing number of teachers working part-time, the total number of teachers who share the task of educating the pupils has considerably increased. As a result of this development the traditional image of a primary school in which a group of children is taught throughout the week by one single teacher is rapidly superseded by an image in which the bemused pupils are confronted with a multitude of different faces. This development has evoked widespread concern with regard to the well-being of the pupils as well as to their level of achievement. This in turn has led to a research being commissioned into the scale of this phenomenon and the effect it has on the pupils.

Method

This paper will focus on the possible negative effects of complete usability of teachers on the well-being and achievement level of their pupils.
In a first step towards studying the possible negative effects, the scale in which the complete usability of teachers occurs was determined in an sample of 800 'basis schools' (the combination of kindergarten and primary school). We received data

on the teachers appointments characteristics from 500 schools, which enabled us to determine the scale of part-time working and complete usability in primary education. The data furthermore allowed for the selection of about 50 schools containing grades 2 (the five year old) and 3 (the six year old) with the right characteristics for a quasi-experiment, aimed at determining the effects on the pupils. In this experiment the schools, for grade 2 and grade 3 separately, were divided into two groups, i.e. a group of schools in which the pupils are taught by two or more teachers who did not possess the adequate training and a group of schools in which the pupils are taught by one teacher with the 'right' training (the control group). In the experiment we focussed on teachers who made use of the regulation concerning the complete usability for only a part of their appointment. We argued that the problem would lie with teachers who had to alternately teach kindergarten children and primary school children and not so much with teachers who teach the 'wrong' pupils for the whole of their appointment. In the first place very few teachers are in this position, and those who do will quickly adapt to this teaching situation.
The two type of schools were compared, by means of multi-level statistical analyses, on the basis of the average language and numeracy skills of the grade 2 and grade 3 children, as well on their level of well-being. The comparison was controlled for the age, ability, ethnic and social economic background of the pupils and for differences in their opportunity to learn (the overlap between the implemented curriculum and the tests used) and the teaching style (e.g. the way the teaching was geared towards individual differences between pupils and the achievement orientation of the teacher).

Results
In comparison with earlier research in this field, this survey indicates a decrease in the level in which schools employ the possibility, offered by the new law introduced in 1985, to let teachers teach pupils for which their professional training do not equip them. In 1988 74 % of the schools questioned, reported the occurrence of complete usability of teachers, compared to 55 % of the schools in our survey held in 1991. In the quasi-experimental research conducted in 50 schools, no significant differences were found between the two types of schools, i.e. schools with and without complete usability of teachers in grades 2 and 3. This lack of effect proved not to be induced by differences in teaching styles.

Conclusion
Despite the often proclaimed negative effects of teachers without the proper training background for the job at hand on the results and well-being of their pupils, no such effect could be found in this study.

Discussion
The results of this study indicate that the quality of primary education is not impeded by the fact that circumstances force school principals to led part of the education be performed by teachers without the proper training. Pupils, as well as teachers, apparently adapt well to this situation.

COORDINATION IN SECONDARY SCHOOLS AND ITS IMPLICATIONS FOR STUDENT ACHIEVEMENT

B. Witziers, Department of Education, University of Twente, Enschede, The Netherlands

Introduction

Regarding school organizations the concept of coordination is subject to theoretical and empirical controversies. One can discern two contrasting views. The first rejects the notion that schools are tightly linked bureaucracies. It accepts the proposition that schools lack close internal coordination and that content and methods of instruction tend to be linked loosely to the control and influence of both the bureaucratic and collegial aspects of schools (cf.Cohen, March & Olsen, 1972; Meyer & Rowan, 1979; Weick, 1979).

The second view states that schools differ on these aspects. Although not as sharply focused on coordination as the literature on 'loose coupling', the growing body of literature on 'effective schools' points to the importance of coordination. Teachers may assume shared responsibility for accomplishing a set of objectives, depending on their shared attitudes toward pupils and shared conceptions of effective instruction. Further, schools may differ in the way school management employs school-wide plans and schedules (structural linkages) with regard to the instructional system and in the way school management controls and evaluates the instructional system (cf. Rosenholtz, 1989).

The above leads to two questions guiding this research project. First, it is clear that teaching is controlled and coordinated, but unknown is the nature and extent of coordination within schools. The first question guiding this research project is related to this problem.

Besides, the effects of instructional coordination are unknown. The reason for coordination is organizational effectiveness, yet much has to be learned about the exact relationship between coordination and organizational effectiveness. This relationship is at the heart of the second question guiding the is study. Main focus in this study is on coordination within departments. Departments are regarded as task groups structurally embedded in the organization of schools. They are responsible for (the coordination of) the instruction of specific subject areas. Research concern ing departments is almost absent. As such, it is a neglected area in the field of educational research (cf. Siskin, 1991). Besides departments, attention is given to the (coordinative) role of school management and to the role of heads of department.

Data source
In this research project 107 departments, representative of both geographic region and school size, in three subject areas (history, mathematics and English) were involved. Data were collected by means of questionnaires. The teacher questionnaire covered questions concerning teacher collaboration, instructional policies, decision-making, team-consensus, cohesion and the intensity and frequency of formal and informal meetings. The head of department questionnaire included questions concerning the school management. These questions referred to the nature and extent of bureaucratic (teacher control and school wide rules governing teacher behaviour) and cultural linkages (school management - teacher communication with regard to the educational process of the department). Besides, questions were asked about the role of school management in the decision-making process. Standardized student scores on a national achievement tests in the field of mathematics were used as achievement scores. Furthermore , data with respect to relevant control variables (e.g. teacher experience, background characteristics of students) were collected.

Methods
Multiple regression analysis was used to establish the relationship between variables. A one-way analysis of variance was employed to explore the differences in the nature and extent of coordination between departments in different subject areas. A path model was developed to determine the influence of different exogenous variables (subject area, role of head of department and school management , school size) on variables concerning coordination within teams. To see whether theoretically relevant patterns of coordination in school organizations could be discerned a cluster analysis was employed. These patterns of coordination were used in a multi-level model to assess the relation between instructional coordination and student achievement. A second multi-level analysis established the relationship between student achievement and individual variables concerning coordination.

Results
The results of this study indicate that school organizations are less loosely coupled than some authors state. Especially the data concerning the way departments coordinate the instructional system lead to this conclusion. Departments employ many procedures and rules to regulate the instructional behaviour of individual teachers. These rules and regulations concern mainly the content of instruct ion, the nature and extent of testing, grading and the goals and outcomes of teaching. The influence of departments on these aspects of teaching in the decision-making process concerning these aspects is also considerable. On the other hand, teacher autonomy prevails within classes. Teachers are autonomous with respect to how to reach these commonly agreed upon goals and outcomes of education . However, this does not imply that within departments teachers differ to a large extent in their classroom behaviour. The data show that within departments educational views and classroom behaviour of teachers are quite similar. The results concerning the differences between departments show that the willingness of teachers to cooperate is by far the most important factor in explaining differences in coordination between departments. Aspects as communication,

instructional policies, collaboration are greatly determined by this factor. The influence of school management can be traced in the extent departments develop policies with respect to their educational process. The head of department influences the intensity and frequency of collaboration an d communication within departments. Furthermore, the data show (large) differences in coordination between departments in different subject areas. This implies that within schools culturally based differences between departments exist. The results of the cluster analysis show that 3 types of school organizations can be distinguished. The first can be characterized b y a lack of close internal coordination. Educational leaderschip, collaboration, communication between teachers and policies concern ing the educational process are (almost) absent. Within the second type the instructional process is mainly coordinated mainly on the basis of collaboration and consensus. Policies concerning the educational process and educational leaderschip are missing. The third type can be typified as tightly coupled or bureaucratic. Main features are educational leadership and strict policies concerning the educational process at the departmental level. Less important features are communication between teachers and teacher collaboration. Results concerning student achievement make clear that schools that can be characterized as tightly coupled systems are more effective than loosely coupled schools. Responsible for this conclusion are mainly the results concerning student achievement and the individual variables concerning coordination. Variables like strong educational policies concerning the educational process and educational leadership have a positive relationship with student achievement. Both aspects refer to the fact that evaluation is possibly the crucial factor in discerning effective from ineffective schools.

References

Abbot, M.G. (1987). Coordination, a neglected variable in the study of schools, in: Johnston, G.S., Yeakey, C.C., Research and Thought in Administrative Theory: Developments In The Field Of Educational Administration.

Abbot, M.G., Caracheo, F. (1988). Power, Authority, and Bureaucracy, in: Boynan, N., Handbook of Research on educational Administration. Longman, New York, 239-257.

Firestone, W.A., Wilson, B.L.(1985). Using bureucratic and cultural linkages to improve instruction: The principals' contribution, Educational Administration Quarterly, 21, 7-30.

Fuller, B., Izu, J.A. (1986). Explaining school cohesion: What shapes the organizational beliefs of teachers?, American Journal of Education, 501-535.

Joyce, B.R., Hersh, R.H., McKibbin, M. (1983). The structure of school improvement. Longman, White Plains, New York, 1983.

Meyer, J.W., Rowan, B.(1978). The structure of educational organizations, in: Meyer, J.W. (Ed.), Environments and organizations. Jossey-Bass, San Fransisco.

Mintzberg, H (1979. The structuring of organizations, Prentice Hall, Englewood Cliffs, New Jersey .

Miskel, C., McDonald, D, Bloom, S.(1983). Structural and expectancy linkages within schools and organizational effectiveness. Educational Administration Quarterly, 10, 49-82.

Rosenholtz, S.J. (1989). Teachers' Workplace: the social organization of schools. New York, Longman.
Thompson, J.D. (1967). Organizations in action. McGraw-Hill, New York, 1967.
Siskin,L.S. (1991). Departments as Different Worlds, Educational Administration Quarterly, 2, 134-160.
Weick, K. (1976), Educational organizations as loosely coupled systems, Administrative Science Quarterly, 21, 1-19.
Wilson, B.L., Corbett, H.D.(1983), Organization and change: The effects of school linkages on the quality of implementation, Educational Administration Quarterly, 19, 84-104.

SCHOOL SIZE, CHARACTERISTICS AND OUTCOMES IN DUTCH SECONDARY EDUCATION

F.G.M. Kleintjes, & E.J.J. Kremers, Cito, Arnhem, The Netherlands

Introduction

In 1988 the Dutch National Institute for Educational Measurement (Cito) carried out a baseline study in the third year of secondary education in the Netherlands. In a sample of 649 schools school- and classroomcharacteristics and performance of pupils were assessed. The results of the study (Kremers, 1990) will be compared with the performance of pupils after the introduction of the so called basic education.

In this paper we present the results of a secondary analysis of the baseline data. We investigate the relations between a number of schoolcharacteristics, classroom characteristics, pupil characteristics and pupil achievement on tests for biology, English, Dutch and mathematics.

Sample and design

The population sampled included pupils from five schooltypes: Those who were taught in the three schooltypes for general secondary education within the Dutch educational system, that is intermediate general education (Mavo), higher general education (Havo) and pre-university education (VWO). Furthermore, the target population consisted of the two largest schooltypes within junior vocational education, that is to say junior technical education (LTO) and junior domestic science education (LHNO). The research population was composed of 20 subpopulations (4 subjects x 5 schooltypes). For each subpopulation a representative sample was taken. In total more than 27,000 pupils form 1,259 classes participated in the baseline study. In order to limit the workload for school and pupils, all pupils participated in the testtaking for one of the four subjects only. The number of items was too numerous te be answered by a pupil within the available testing time. (Dutch: 180 items; English: 280; Biology: 280; Mathematics: 260). We decided therefore to choose a matrix sampling design, in which the items of each schoolsubject are divided over a number of test booklets according to the principle of 'nominal equivalence'. This required that we had to create comparable test booklets on the basis of content specifications. The tests were the same for pupils of the various schooltypes and they therefore contained questions of diverse difficulty. The allocation of booklets to pupils was at random.

Method

In order to distinguish between the effects of school, classroom, and pupil characteristics, that were available for analysis, a multilevel approach was opted for. Our main interest was to explore the effect of school level characteristics on pupil achievement. Main effects of classroom and pupil characteristics had already been highlighted in earlier reports such as in the one mentioned above. The latter characteristics have been included in the present research in order to give more insight in relations between the employed variables at the various levels. The available variables at the three levels have been put into a main effects model. Pupil level variables included were: Ses, sex, number of years spent in secondary education and ethnicity.

One classroom variable was available: opportunity to learn. The school variables used were: kind of school (categorical or combination of schooltypes such as f.g. LTO/Mavo, Mavo/HAVO/VWO), denomination, total size of the schoolpopulation and the number of inhabitants per square kilometer of area of the school.

Results

Pupil, class and school level variances have been fitted seperately for each schooltype. These variances differ quite a bit among schooltypes and subjects, indicating the variety in achievement of students over and in schooltypes. A model with main effects only, taking into account the available pupil, class and school variables, reduces the variance at school level to about half the amount of a base line model, showing an excellent incremental fit. The variance explained ranges from fifteen percent to almost all for the different schooltypes and subjects under consideration. An effect is to be interpreted as a difference in means of the populations that are considered, after correcting for differences due to background variables. A 'pure' effect, so to say. An attractive feature of using a multilevel approach is that the effectsize, defined as the ratio of the effect and the within group variance, can be evaluated at both student and school level. An effectsize of 0.2 is usually considered to be small, one of 0.5 medium and one of 0.8 large. Since the residual variances vary over schooltypes and subjects, effectsizes do as well. In this way it has been possible to present a detailed state of affairs by schooltype. Major findings are that the effect of the size of the school did not turn out to be significant, with the exception of an effect of schoolsize 360-500 for Mathematics, the effectsize being 0.30. This means school size does not explain difference in pupil achievement, only for mathematics pupils in a school with size 360-500 do a bit better.

It was found that categorical schools do have an effect, its size was 0.50 for Mathematics and 0.60 for Biology; no effects for Dutch and English. A variety of combinations of schooltypes is found in the dutch schoolsystem, taking these up as explanatory variables in the model did not result in a significant effect.

It means that being in a categorical school has a positive effect on pupil achievement. Corrected for for all available background variables those pupils do better on biology and mathematics. Being a pupil in any other kind of school (such as combined LTO/Mavo, Mavo/Havo/VWO etc.) does not have an significant effect on achievement. In the model employed, the effectsize of denomination Protestant for Mathematics and Dutch was 0.20. No effect was found for English and Biology. The number of inhabitants (in hundreds per square kilometer) in the area of the

school had a negative effect with a size of 0.10 to 0.20 for three of the four subjects, only for English no significant effect was found.

Conclusion and discussion

So far 'only' a main effects model has been considered. However, even from this, it is clear that our findings are relevant for policy makers as well as for researchers. In the paper presentation the results will be discussed regarding the plans the Dutch government has proposed for increasing school size and the establisment of combined schooltypes ('comprehensive schools') in secondary education. More complex models including interactions at various levels are explored meanwhile. The results of that analysis are to be presented at the conference as well.

ECONOMICS OF SCALE AND SCOPE IN SCHOOLS: AN ECONOMIC ANALYSIS BY MEANS OF TRANSLOG COST FUNCTIONS

Marc Vos, Higher Institute of Labour, Catholic University, Leuven, Belgium

This study examines the efficiency, from a micro-economic perspective, of 209 Flemish educational establishments providing primary and/or secondary education. The costs analysed relate to staff and operational costs financed by the Education Department. School output is actualized by means of variables which reflect the size of the relevant pupil categories. Mathematical cost functions are estimated using the regression analysis technique. The basic function is a three-product translog cost function. This working method implies a tacit application of the "efficient production technology" hypothesis, though this is mitigated to some extent by the introduction of institutional and other heterogeneity variables. The estimates produce interesting findings relating among other things to the number of school locations, synergetic effects and effects of scale.

The evaluation of educational facilities takes place largely on the basis of the effectiveness criterion: schools are effective if they adequately achieve the predetermined objectives (increasing the level of human capital, socialization, instilling discipline, democratization, etc.). On the basis of this effectiveness criterion alone, however, the contribution of schools to the welfare of society cannot be evaluated to a sufficient extent. Schools, after all, operate using "scarce" manpower and material resources, on which they maacty be able to bestow to some extent an alternative, more productive direction. It follows from this that educational facilities can only make an optimum contribution to social welfare if they achieve their objectives (at minimum costs). This efficiency problem has been actualized from an economic perspective as follows. Educational establishments have in common with economic institutions, such as commercial enterprises, that they transform production factors (staff and material resources) and raw materials (inflow of pupils to be taught) into outputs (outflow of educated pupils). Educational facilities optimize their contribution to the welfare of society if:

- all schools have both a technically and an economically efficient production process;
- the predetermined output is distributed among schools in such a way that each school attains its optimum size. This optimum size is achieved in accordance with the level and structure of the production, with average costs being minimized.

In this study, school efficiency was examined from a micro-economic perspective using a sample of 209 Flemish educational establishments providing primary (3-12 years) and/or secondary (12-18 years) education.

The starting point for this study was a descriptive analysis which enabled the determinant variables to be distilled from the cost price. The school costs analysed relate on the one hand to staff costs, and on the other to the so-called operational costs. The latter category contains the variable costs (exclusive of staff costs) and a proportion of the fixed costs. Only those costs which are financed by government grant were taken into account. School output was actualized on the basis of variables which reflect the size of the relevant pupil categories. The large number of potential determinant variables and their complex interrelationships, however, make it necessary to supplement such a descriptive analysis with an econometric study. In the present study it was therefore decided to estimate the mathematical cost functions using regression analysis techniques. The basic function estimated was a three-product translog cost function, in which the natural logarithm of the total operating cost is a function of the natural logarithm of the pupil numbers in general
secondary, technical and vocational education and in primary education. Such a specification has attractive aspects and makes possible interesting applications:

- calculation of synergy effects;
- price homogeneity;
- flexibility with respect to scale and cost elasticity and in respect of substitution elasticity;
- simple calculation of all kinds of criteria.

This working method does, however, imply tacit application of the "efficient production technology" hypothesis. As a result, the analysis of the efficiency problem is incorrectly reduced to a determination of the optimum size and structure of the school output. However, by basing the econometric study on situational efficient production technologies through the introduction of institutional and other heterogeneity variables, this hypothesis is somewhat mitigated.

The econometric analysis of the operational costs based on the translog specification produces several interesting findings relating to Flemish educational establishments. Thus, for example, the number of locations covered by a primary school proves to have an essential influence on the level of the average operating costs. From an economic standpoint, those primary schools which restrict themselves to a single location, and to a relatively small size, function most efficiently. If a primary school has exceeded its optimum pupil level (from the point of view of the operating costs), then the loss in efficiency which is associated with the growth of the school can be offset to some extent by opening sub-schools. The greater the number of locations covered by a primary school, the higher is the optimum pupil number. Technical and vocational secondary schools are relatively expensive due to their principal orientation towards technical and practical subjects. As the complexity increases, (reflected in a reduction in the number of pupils per subject), these schools also become more expensive. These schools do, however, appear to be able to achieve a higher optimum pupil number than general secondary schools. The chance of increasing scale benefits in technical and vocational education increases if the increase in the pupil number is achieved through an increase in the number of pupils per subject. Offering primary and

general secondary education together, combined with the provision of technical and vocational secondary education, leads to important synergetic disadvantages. However, these can be offset to a considerable extent through adequate decentralization of the finance and resources policy.
The econometric analysis of staff costs on the basis of the translog specification shows in the first place that there are continuously increasing benefits of scale, with the negative effect of pupil numbers on the average staff costs increasing as the total staff costs increase. This negative effect has the greatest impact in general secondary schools.

EFFICIENCY OF REGIONAL AND LOCAL SYSTEMS OF EDUCATION

A. Petrikás, Department of Education, Lajos Kossuth University, Debrecen, Hungary

Introduction

When interpreting the theme we started out from the principle that creating the system of education is the basis of the strategy of general education and development. The individual mosaic-like educational efforts - either as issues of content or method - cannot be successful in the development of educational qualities. The reasons are as follow:

- The effects isolated from one another do not ensure the individual's optimal development.
- The effects may solve certain educational situations but do not ensure the link between education and self-education (Á. Petrikás, 1983 and 1987, 1990; Z. Báthory, 1985; G. Halász, 1990). To interprete the educational process as a system assists in institutional harmonizing, organic shaping of social claims and individual requirements.

Conceptually the notion of local educational system is to be meant by the effective activity of an educational institution or group of institutions at the individual's residence according to pedagogical concept and specific system of requirements (Á. Petrikás, 1983 and 1991) Stimuli, of the local society, traditions, pedagogical forces of the society and institution create the necessary system of effects.

A regional system is a natural combination of the local ones within which self-guidance and self-organization operate decision-making mechanisms.

In our hypotheses the alternative variations, creation of an individual aspect make these systems suitable for organic personality-development. Thus development and research serve this purpose.

Methodologically the main results were ensured by case-investigations:

- action analyses;
- survey of mass data, 1989;
- evaluations of individual variations;
- comparison of process modules.

Results

Our results based on research conducted (at about 30 educational institutions) between 1980-1990 (Á. Petrikás, 1991) are as follow:

- The regional and local educational systems become efficient, if variants are established and the unity of development and self-development is ensured.
- The appearance and development of dominant activities reflect the state of the pedagogical concept.
- The individual aspect is created by an enriching value- system, specific norms, moral, psychological and mental cohesion.
- Self-developing, alternative pedagogical programmes prove expanding efficiency (place of residence, inter-institutional relationship, establishment of regional situation).
- Chances for self-guidance and self-governing and their state are specific indicators of efficiency.
- Under our conditions the durability of effects is ensured by democratization, human relations and inter-personality connections.

Conclusion

As a summary it can be stated that according to our hypotheses the local and regional educational systems reflect the principles of organic development. To intensify their efficiency highlights the following factors:

- decentralization of guidance;
- self-governing educational institutions;
- the necessity of differentiated educational strategy (on micro-, mezo- and macro-levels) for ensuring the individual's optimal education.

Discussion

It is worth discussing how to overcome factors, social controversies inhibiting the activity of self-developing educational systems. To what extent are the frames of an autonomous school, kindergarden, educational workshop to be shaped? How to battle against negative effects, loneliness, formalism and pedagogical manipulation?

Debates, comparisons with European experiences may help solve these complex social-pedagogical problems

References

Petrikás Á. (1983) A nevelő közösség funkciói és az iskolai nevelőmunka rendszerének fejlesztési stratégiái Akadémiai doktori értekezés, és tézisei. Debrecen. Functions of educational systems and strategies of development, of pedagogical work at schools. Academic doctoral dissertation and theses.

Petrikás Á. (1987). Közösség és személyiség az iskolai nevelés folyamatában. (társszerzővel) Közösség, személyiség, alternativitas, döntés az iskolai nevelésben Tankönyvki ado, Budapest. Community and personality in the pedagogical process at schools. (with co-author) Community, personality, alternatively, decision in education at schools.

Petrikás Á. (1990). Egyén, közösség, nevelés Elméletek az iskolai egyénnevelésről II, pp. 3-20. OPKM, Budapest. Individual, community, education. Hypotheses of educating individuals at schools II.

Báthory Z. (1985). Tanitas és tanulás Tankönyvkiadó. Teaching and learning. Budapest.
Halász G. (1990). Iskola, helyi társadalom, iskolatanács. Akadémiai Kiadó, Budapest. Schools, local societies, school-councils.
Petrikás Á. (1991).Önfejlesztő pedagógusok - önkormeanyzó iskoleak. Debrecen, KLTE. Self-developing teachers - self-governing schools.

EFFECTIVE INSTRUCTION AND GROUPING OF STUDENTS

G.J. Reezigt, & M.G. Weide, Institute for Educational Research (RION), University of Groningen, The Netherlands

Introduction

One of the major problems in educational research is the separate development of various lines of research. An example is the research in the field of effective instruction and the research in the field of grouping students. Although reviews in both fields of research nearly always indicate some links, there are hardly any empirical studies combining ideas from the literature on effective instruction and on grouping.

In research literature on effective instruction, grouping is always mentioned as one of the important variables in creating an effective learning environment for students (Levine & Lezotte, 1990; Creemers, 1991). However, it seems to be virtually impossible to decide which way of grouping is most effective. A reason for this are the contradictory results studies on the effects of grouping on student achievement have yielded. Another reason is the lack of studies on grouping that also pay attention to instructional variables.

Research on grouping typically ignores the way grouping is implemented in classrooms and the way instructional variables may influence this process. Grouping is nearly always treated as an isolated variable. The suggestion is implicitly made that teachers all show the same teaching behaviour and differ only in the way they group their students for instruction. Only fairly recently some authors (Slavin, 1987a, 1987b, 1991; Gamoran, 1986; Gamoran 1991; Gamoran & Berends, 1987; Gamoran, Berends & Nystrand, 1990) suggested that this very situation may account for the contradictory results, as very little is known about the actual implementation of grouping. Is is a well-known fact, however, that many teachers meet serious obstacles when implementing any form of grouping (Eder, 1981; Arlin, 1982); and that teachers who make use of the same kind of grouping may be very different.

The few studies explicitly looking at grouping characteristics as well as at instructional characteristics of teachers clearly show the surplus value of the combination of both groups of variables in explaining effects in student achievement (Gamoran, 1984, 1986). An important advantage is the possibility to estimate the effects of grouping more accurately.

This paper will try to answer the question whether effects of grouping on student achievement are mediated by instructional characteristics of teachers, and to what extent.

Method, data source, techniques

In 1987 and 1988 data were collected in a random sample of 220 elementary schools in the Netherlands. In 1987 students and teachers of grades 5 and 7 (age groups 9 and 11) cooperated, in 1988 the same students and their teachers of grades 6 and 8 (age groups 10 and 12). In total 10,000 students and 1000 teachers participated by filling in written questionnaires and taking standardized tests.

At teacher level information is available on the way students are grouped (four different forms, all found in Dutch elementary education, are distinguished: whole class instruction, grouped-based mastery learning, within-class ability grouping and individualized instruction). In addition to this, information is available on several indicators of effective instruction (emphasis on basic skills, homework, frequent assessment of student progress and monitoring). Data are also available on experience as a teacher and on some class characteristics (class size, mixed age class vs. single age class, degree of actual achievement heterogenity).

At student level intelligence, achievement in language and math were measured. As the data concern two levels of research, multi-level analysis was used by means of the VARCL program for variance component analysis. Student achievement in 1987 is a covariate, while student achievement in 1988 is the dependent variable.

Results

The effect sizes of grouping (compared with whole class instruction) vary from nearly zero for mastery learning to - .60 for ability grouping and individualized instruction. When prior achievement is taken into account, the strong negative effects of ability grouping and individualized instruction diminish somewhat.

The proportion of variance at teacher level is about 20 percent. Student variables explain most of the variance; when they enter the analysis, only 7 percent of variance at teacher level remains. This means student characteristics are by far the most powerful predictor of achievement.

After controlling for student characteristics, instructional characteristics explain a larger proportion of the remaining teacher level variance than grouping procedures (25 vs. 10 percent). When instructional characteristics enter the analysis, the effects of grouping are affected sometimes (but not always). However, the changes are not consistent: sometimes the effects of grouping become more negative and sometimes they become less negative or even insignificant. The instructional variables responsible for the changes vary from one grade to another.

Although the results of the analysis do not show a consistent picture, they make clear that research on effects of an isolated aspect of teaching (in this case, grouping) is not very informative. This is even more so when corrections based on prior achievement can not be made, as this leads to an inflation of effects. Several indicators of effective instruction should always be part of research that pretends to explain achievement differences between children, and research in this field should always include at least two measurements of student achievement.

References

Arlin, M. (1982). Teacher responses to student time differences in mastery learning. American Journal of Education, 90(4), 334-352

Creemers, B.P.M. (1991). Effectieve instructie.'s-Gravenhage, SVO.
Eder, D. (1981). Ability grouping as a selffulfulling prophecy: a micro-analysis of teacher - student interaction. Sociology of Education, 54(3), 151-162.
Gamoran, A. (1984). Egalitarian versus elitist use of ability grouping. New Orleans, microfiche ED 245821.
Gamoran, A. (1986). Instructional and institutional effects of ability grouping. Sociology of Education, 59(3), 185-198.
Gamoran, A. (1991). Alternative uses of ability grouping in secondary schools. AERA-paper, Chicago.
Gamoran, A., Berends, M. (1987). The effects of stratification in secondary schools: synthesis of survey and ethnographic research. Review of Educational Research, 57(4), 415-435.
Gamoran, A., Berends, M., Nystrand, M. (1990). Classroom instruction and the effects of ability grouping: a structural model. AERA-paper, Boston.
Levine, D.U., Lezotte, L.W. (1990). Unusually effective schools: a review and analysis of research and practice. The National Centre for Effective School Research and Development.
Slavin, R.E. (1987a). Mastery learning reconsidered. Review of Educational Research, 57(2), 175-213
Slavin, R.E. (1987b). Ability grouping and student achievement in elementary schools: a best evidence synthesis. Review of Educational Research, 57(3), 293 - 336
Slavin, R.E. (1991). Ability grouping in the middle grades: effects and alternatives. AERA-paper, Chicago.

INFLUENCING EDUCATIONAL PRACTICE THROUGH PERFORMANCE INDICATORS

P.B. Tymms, Moray House Institute of Education of Heriot-Watt University, Edinburgh, Great Britain

The problem

School effectiveness research has provided us with the knowledge and techniques to produce fair performance indicators. But does their use improve educational provision? Are fears of league tables justified? How do teachers who have actually been part of a system which rank orders their performance actually feel about the system. Do teachers change their teaching in response to feedback and if so what kind of feedback is the most appropriate?

ALIS has been feeding back quantitative summaries of educational performance to schools for nine years and qualitative evaluations suggested that:

- Complex numerical reports were an obstacle to the acceptance of performance indicators by some teachers.
- In-service provision was a major factor in determining teacher's involvement with the data.

Could hard evidence be found to support these hypotheses?

Method

A questionnaire was sent to 804 teachers involved in ALIS and the results of the survey were used to assess the impact of the performance indicators and to compare the responses of various categories of respondent.

The data were also used to evaluate the results of two interventions. In the first of these two forms of one of the three reports sent to schools and colleges were randomly assigned to departments within schools and colleges. One of the forms was detailed and included several tables of data whereas the second form was short. It contained no tables and simply summarising the findings of the statistical analysis.

In the second intervention teachers from randomly chosen departments were invited to attend ALIS workshops on Saturday mornings.

These experiments were designed to test two hypotheses about involvement with the system derived from evaluation exercises. The first being that simplified reports would have more impact than more complex reports and the second that in-service would be important in influencing attitudes and use of performance indicators.

Result

Analysis, using multi-level models, indicated differences between schools and between Local education Authorities (LEAs) in the way in which ALIS was used and appreciated. Estimates of about 7% of the variance of involvement at the LEA level and about 15% at the institution level would seem to be reasonable.

At the teacher level it would appear that about half the A level teachers involved in ALIS had read the reports and discussed then within their institutions. About a fifth of teachers had changed their teaching in some way as a result of ALIS.

Generally the teachers found ALIS to be little trouble, to regard it as fair, quite useful, value for money but less than friendly.

Heads of Department (HoDs) tended to read the reports more than others, as did those who had been to in-service workshops, although the HoDs' attitude towards ALIS was similar to members of their departments.

Those who had been on in-service tended to use ALIS and to rate it as being more valuable and more needed than others. The analysis implied that in-service attendance was crucial so far as use of and attitude towards ALIS is concerned.

There was no evidence that the number of years teaching A level had any relation to the use or perception of ALIS, but the longer that ALIS had been known about the more use it had been put to, and the more positive were the responses to it.

Language teachers were less positive about ALIS than Maths/Science and Humanities teachers.

Results from the First Intervention

Not surprisingly the short report was generally read more than the long report and the indication was that the type of report had little or no effect on Maths/science respondents but that within the Humanities area long reports were associated with more discussion and more changes to teaching changes.

On the other hand Language respondents were more favourably disposed towards ALIS when they had received short reports. They were also more likely to say that they made changes to their teaching practice as a result of ALIS.

Results from the Second Intervention

Those who had been invited to attend workshops were more likely to have read the reports and to report having discussed matters relating to ALIS and having changed their practice than those who had not.

There was also evidence that Saturday morning invitations had a beneficial influence on the attitudes of Language and Humanities specialists towards ALIS in that it was perceived as more friendly, more useful and more valuable by those who were invited.

Conclusion and discussion

Whilst teachers varied in their responses the data indicated that it is perfectly possible to run a performance indicator system in which comparisons are made in league tables and in which teachers are appreciative of the information provided.

No doubt there are many features of ALIS which have made it an acceptable system but the results indicated that communication through appropriate in-service and reports is particularly important.

The analysis also indicated differences in the responses of teachers in different curriculum areas. It seemed that Language teachers were most responsive to report simplification and to in-service provision and that maths and science teachers were less likely to be influenced. However, those same maths and science teachers were generally more positive than Language teachers. Humanities teachers also responded positively to in-service provision but appeared to be more stimulated by complex reports.

As performance indicators assume a greater importance within education is becomes more and more important to investigate the ways in which performance can be monitored in an acceptable and defensible manner. This paper has indicated that performance monitoring can be acceptable. It has also shown how an indicator system can improve its own provision not simply by surveying the recipients of the data but by experimenting with different forms of provision and monitoring the results - indicator systems can test themselves.

FROM SCHOOL EFFECTIVENESS TO SCHOOL IMPROVEMENT: SOME OBSERVATIONS ON THE STATE OF PLAY

John Gray, QQSE Research Group, Education, Sheffield University, United Kingdom

After a decade of research effort there is growing interest in incorporating the findings of school effectiveness research into strategies for school improvement. Indeed, there is already some evidence on which to draw. But there is also a deepening awareness of the problems to be overcome.

Both research traditions have generated important findings. School effectiveness researchers can show how much schools make a difference; they have developed more appropriate frameworks for analysing schools' performance; and they have identified some of the factors that have/have not been associated with differences in effectiveness. School improvement researchers, in turn, have developed evidence about the processes of school reform, showing how planning and staff participation have been central to the generation of the commitments change efforts require. Change programmes premised on the findings of school effectiveness research have, however, been a good deal less frequent and, in some cases, less successful than might have been expected. Have the problems been related to the knowledge-bases, the planning and activities that seem to flow from them or the difficulties that emerge in actually implementing changes?

Four propositions offer a useful framework for exploring the available evidence. These are:

- that school effectiveness research has been insufficiently focused on identifying strategies that will bring about school improvement; the central findings have not yet properly identi-fied the factors that can be readily changed and developed.
- that the school improvement literature has been insufficiently focused on analysing strategies for securing changes in relation to the outcome measures which have been of prime interest to school effectiveness researchers.
- that key actors (both within and outside schools) have been relatively poor at diagnosing the strengths and weaknesses of individual institutions (especially ineffective ones) and at framing appropriate change strategies in relation to them.
- that implementing changes in ineffective institutions is a good deal more difficult than researchers and practitioners have hoped; most of the factors which underpin an institution's `performance' are best described as `resistant to change'.

All four propositions can be justified to a greater or lesser extent in terms of the available evidence. The central problem remains that of persuading change

agents and practitioners to 'own' the school effectiveness paradigm. To achieve this there must also be changes in the central questions it addresses.

SCHOOL AUTONOMY, PERSONNEL POLICY AND EDUCATIONAL INNOVATION

M. Vermeulen, H.J. Visser, Q.H.J.M. van Ojen, & P.N. Karstanje, Centre for Educational Research, University of Amsterdam, Amsterdam, The Netherlands

Introduction

The Dutch Department of Education seeks to enlarge the autonomy for schools, to enable them in developing a school policy of their own. An important instrument to effectuate this aim, is to provide schools with a budget to spend at their own discretion. This so-called 'lump sum system' however, has not yet been implemented in schools for primary and secondary education. In stead of a lump sum system for these types of schools a kind of transitional measure will be introduced in order to facilitate the chance from strict central legislation to the relative freedom from the schools in a lump sum system. This new system is called the "Formatiebudgetsystem" (FBS) and its application will be restricted to the personnel policy within schools.

One of the reasons for the Dutch Department of Education to implement such a staff budget system was the lack of capacity of policy making on matters of staff development and personnel. According to the department, within the FBS system, school boards will have the opportunity creating a personnel policy optimal fitting for their situation and their goals.

As a consequence, schools ought to be able to relate personnel policy to priorities and needs arising from educational developments. One of the assumptions is that there is a relation between the way schools develop their educational organisation or innovation, and the way they are managing personnel policy and it is presumed that organisational development is related to a certain level of educational innovation. This assumption was one of the objects of the research project[1] The central question of this paper is: "Will more autonomy on school level in personnel policy making, stimulate educational innovation?"

The theoretical concept of the policy making capacity

From the literature on school organisation and school policy making, we expected that schools with a high level of educational innovation like internal differentiation in the classroom, a pupil counselling system, working on a school policy plan, should have made to some extent organisational arrangements in the school in order to adapt, implement and continue these innovations. This theoretical concept is adopted from Ernst Marx (1987;1990)[2]. This theory focuses especially on the

way schools use the degrees of freedom. The policy making capacity of schools is conceived as the mechanism for using the available space for policy making. For example, some schools play everything strictly according the regulation, whereas on the other hand some schools seeking to find every opportunity to use regulation for their own purpose.
Marx distinguishes three major types of schools, according to the extent of their policy making capacity. Several aspects are important in this theory on policy making capacity: the existence of a school mission, the kind of school culture, the role of the school management, the school organisation, the innovativeness of the school.
These three major types are briefly characterized.

Administrative capacity

- Policy making is dominated by financial, legal status, administrational affairs.
- Teachers are mainly vocational experts and have a great autonomy within their classrooms.
- The school culture is individual culture.
- The organisation is a segmental model. Management and teachers function in separated zones. - There is no school mission supported by the whole staff.

Operational capacity

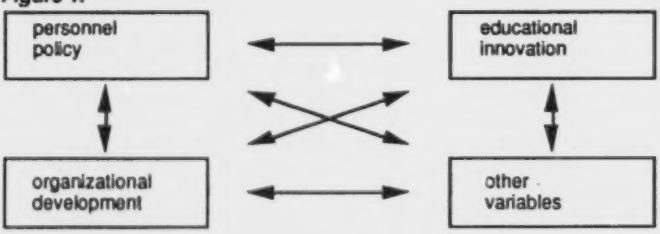
- Policy making is characterized by a relationship between the administrative educational and organisational arrangement in the school.
- Teachers cooperate especially on pedagogic and didactic topics, they have less autonomy in the classroom.
- The school organisation is characterized through a collegial model, with different departments. These departments are relatively autonomous.
- Within the departments people cooperating and consulting each other.

Innovation capacity

- In this type of school they have strategic policy. This means that there is a consistence in policy making within the school, and also in relation with the conditions outside the school.
- The school has a clear school mission.
- The school organisation is flexible, through smaller units who are relatively autonomous. It is a modular model.
- The school is able to assimilate a lot of information, and to give feedback to the participants about their functioning.
- Teachers have many roles. Cooperation is necessary on many topics and classroom autonomy hardly exists.

These types of policy making capacity are supposed to be cumulative, which means that the aspects of a higher level of policy-making capacity include the aspects of the lower ones. This theory enabled us to develop an instrument to determinate the position of schools, on a capacity-continua. According to this theoretical framework, the assumption is made that personnel policy is related to educational organisation and innovation in the school. In other words, schools using fully possibilities of the FBS should by characterized by type 3 of the Marx construct. To test these hypotheses a sample with a fraction of 0.25 was taken random from the whole population of Dutch secondary schools. By means of a survey we tried to find out how schools develop personnel policy, how the school policy can be characterized, what kind of educational innovations they have implemented and what kind of school organisation they have. The response on the survey was 44%. By means of statistical analyses we investigated the relations between the different aspects as shown in figure 1.

Figure 1.



Empirical evidence

From the data there was no evidence that schools had developed a coherent personnel policy. Isolated aspects of staff development were found. For example some schools have structural planned dialogues between the management and teacher about the functioning of the teacher. However no relation is found between these dialogues and for instance formalization of in service training or special support for new teachers.

In short, the data shows isolated aspects of personnel policy in schools, but in most schools it is not yet developed to a consistent policy.

The same trend has been found for educational innovation and for the development of the school organisation. School innovations mostly stand alone, are not related to each another nor to the organisational structure of the schools.

Nevertheless, according to the schools, the decisions how to employ the personnel or to appoint personnel to available functions are based on educational or organisational motives. Apart from this we couldn't find any significant relation between personnel policy and educational innovation, or the development of school organisation.

Only some of the situational or context variables turned out to have significant relation with one of the topics in our research.

The larger schools are, the more they show activities in matters related to personnel or staff development. In such schools we found more often a formal training scheme, a policy plan and terms of references for the various employees.

A relation was found between school complexity (schools with more streams) and the presence of pupil counselling or a pupils guidance system.
An exception must be made for vocational education (lbo-stream). These schools all have pupil counselling and internal differentiation.
An other relation is found between extra facilities for innovations and educational innovations. This relation however, was predictable since most of these subsidized projects, stimulate educational innovations like differentiation within the classroom, pupil counselling and guidance.

Conclusion

Hardly any significant relation is found between variables *within* the various topics and *between* the various topics of our research.
This could indicate that developments in schools are relatively separated processes initiated more or less by coincidence. Naturally there are schools where the supposed relations can be found. Their amount however is to small to have any influence on our general results.
A possible explanation for the absence of relations between personnel policy, school organisation and educational innovation is that schools are able to develop educational innovations, without the need to arrange systematically the supposed conditions in the personnel or organisational area. This idea joins the theory of school organisations as loosely coupled systems (Weick 1976)[3]. An alternative explanation is that relations can't be found because the development in schools are in generally to weak, or in a beginning stage.
Our findings can't confirm the assumption made by the department of education, that autonomy stimulates educational innovations. However it is not rejected, autonomy could be a condition for innovation processes.

References

1. Ojen, Q.HJ.M. van, M. Vermeulen, H.J. Visser, P.N. Karstanje. Personeelsbeleid, onderwijsvernieuwing en het formatiebudgetsysteem. SVO-project 9122. A.B.C. de Lier, 1992.
2. Marx, E.C.H. In: Genemans, L. Autonomie van scholen en deregulering. Verslag studieconferentie, Nijmegen 6-2-1987. Marx, E.C.H., R. Elte, Autonomiseringsprocessen in het onderwijs: Scholen en middenveld-organisaties. Leiden, 1990.
3. Weick, K.E., Educational Organizations as loosely coupled Systems. In: Adm. science Quarterly, 21, 1-19, 1976.

CONTEXTUAL INFLUENCES ON SCHOOL-EFFECTIVENESS: THE ROLE OF SCHOOL BOARDS

Roelande H. Hofman, RION, Institute for Educational Research, University of Groningen, Groningen, The Netherlands

Abstract

This paper is based on a research report that provides information about the influence of school boards, their characteristics and their administrative control, on the effectiveness of the schools they rule. A combination of two lines of inquiry, schooleffectiveness and organizational effectiveness, lays the foundation of the framework of school boards' administrative control used in this study. In this model the school boards are part of the context of schools. An aselect sample of 150 school boards and one specific primary school they rule, was drawn. The data used in this study have a hierarchical structure and the results show that the four clusters of variables from the model do explain variance in cognitive achievement. From the results it can be derived that school boards do make a difference. Some school boards with specific characteristics manage schools with relatively better results in the cognitive domain.

Introduction

Only a few studies examine the influence of the amount and the quality of administrative control on the effectiveness of schools. This is remarkable, since the governing authorities at the level above the school are legaly responsible for the quality of the schools they rule. In this article we report our findings in a study carried out on the role of school boards, their characteristics and their administrative control, on the effectiveness of the schools they rule.

Research questions

- What types of administrative control do schoolboards develop and which characteristics of school boards can be ascribed to these differences?
- Can characteristics of school boards and the administrative control they develop explain variance among schools in pupils achievement?

Theoretical framework

In order to investigate the the two research questions listed above, two lines of inquiry were combined to develop a conceptual model for school boards administrative control. The first line of inquiry focusses on the results of research

on effective schools. Scheerens (1990) developed a socalled 'integrated model of school effectiveness', an context-input-proces-output model. This model lays the foundation of the framework of school boards' administrative control used in this study. In this model the school boards are part of the context of schools. School boards can enhance or diminish the effectiveness of the school they rule, they can have a contextual influence on the functioning of schools. The second line of inquiry focusses on the results of research on organizational effectiveness in general, more specific research on administrative control of school boards. In this line of research the work of Mintzberg (1979) on organizational management is used as a basisc for describing organizations. Mintzberg (1979, 1983, 1989) uses different types of variables to describe an organization. Important variables are first of all the influence of the institutional environment on the structure and functioning of an organization, which he calls the contingency factors. For the description of the internal structure of an organization he uses so-called design parameters. The combination of this two lines of inquiry is necessary to expand knowledge on which activities of school boards can enhance or diminishe the quality of the level below, the schoollevel (Purkey & Smith, 1983; Bidwell and Kasarda, 1975).

Methodology

Data source and instrumentation

An (aselect) sample of 146 school boards was drawn out of a total of 8000 Dutch primary schools. Data collection methods used were questionnaires for school boards and for schoolleaders, and the testing of pupils in groups 6 and 8 of the primary schools.

Operationalization of variables

At the context-level, the characteristics of school boards were measured after Mintzberg (1989) as contingency variables and as organizational design parameters of school boards (alpha coefficients of the constructed scales lie around .80). The activities of administrative control that school boards develop were measered on six dimensions (reliable scales _'s from .72 tc .86).
At the input-level we take into account differences between schools in resources and in characteristics of the schoolteam and of the school environment.
At the schoolproces-level six indicators of schooleffectiveness have been conctructed: educational leadership; degree of achievement oriented policy; evaluation of education on the schoollevel; emphasis on basic skills; orderly and formally stated school atmosphere; degree of pupil-evaluation oriented policy. The constructed scales are highly reliable with alphas around .80 or are strong (hierarchical) Mokkenscales H-coefficient = >.53).
At the output-level pupil attainment scores in language and in arithmetic were measured with standardized achievement tests. By means of the VARCL multi-level program (Longford, 1986) schooleffectiveness-indices per school (gainscores) were constructed. These indices use pupilscores for arithmetic and language achievement and these are controlled for pupil characteristics as previous achievement, sex, ses, intelligence and ethnicity.

Method of analysis

The data used in this study have a hierarchical structure, so it seems appropriate to work from an 'inside-out-regression-analysis' (see Creemers & Scheerens, 1989). The clusters of relevant variables were brought into a step by step regression-analysis with at the first step the covariates, at the second step the input-variables, at the third step the schooleffectiveness process indicators and at the last step the context-effect of school boards containing the contingency variables and the organizational design parameters of the school boards.

Conclusion and discussion

- The activities of school boards are more likely to be on the areas of management like on staffing and on financial matters and less on the educational matters.
- The contingency variables and also design parameters do very much ascribe to differences in the school boards' activities.
- Multi-level analyses show that the percent of between school variance that is connected with differences among schools is 12% for arithmetic and 8% for achievement in language.
- 42% of the between-school variance in arithmetic and 20% of laguage can be explained by the selected variables (covariates, input, proces, context).
- From the results it can be derived that school boards do make a difference. A chacteristic of the functioning of the school board that is related to the effectiveness of schools is the influence of members of the schoolorganization on the decisions of the school board. This finding can be of importance for the discussion on the governing structures in general; individual school boards could learn from it that by changing their decisionmaking proces they can make a more fruitful contribution to the effectiveness of the schools they rule.

EXTERNAL SUPPORT, SCHOOL LEADERSHIP AND TEACHER BEHAVIOUR IN ELEMENTARY EDUCATION: A MULTILEVEL STUDY

K.M. Stokking, Department of Education and Educational Research, University of Utrecht, Utrecht, The Netherlands

Research questions and context

Various investigations of primary education carried out in the Netherlands in the 1980s show that there are differences between schools. In addition differences between regions were found. Research into differences between classes within schools is rare. Hardly any use has been made so far in this area of the recently developed methods and techniques of multilevel analysis.

The law, as well as government policy, aims at innovations in education by way of individualization, differentiation and extra care (these three aspects will be referred to as IDE). Research into IDE has produced conflicting results (Stokking 1992).

In the years 1989-1990 a multi-level investigation was carried out into the following questions: (i) with regard to IDE, how great are the differences in primary education between classes and between schools, and (ii) what are the factors at the levels of class (teacher), school and region by which these differences can explained?

This study, itself unfinanced, was a spin-off of several research projects which were financed by the Institute for Educational Research in the Netherlands (SVO). These projects studied respectively the utilization effects in primary and secondary schools after dissemination of unasked-for information (grant ANGO 1423) and the research and policy in the area of IDE during 1984-1990 (grant ANGO 0335). This research is part of a larger research programme on the dissemination and utilization of information in the Dutch educational system. In most projects in this programme the topic of IDE (a didactic innovation) is chosen as the content of the information of which the dissemination and utilization is studied. In continuation of this particular study a project is now started, again with a grant from SVO (ANGO), in which the data will be analysed in a number of different ways, which all have been used or suggested and/or are currently being used for the analysis of a multilevel problem.

Relevant theories and choice of variables

The choice of the domain of IDE was not only based on the actual national innovation policy but also on the fact that the needs of teachers for inservice training are the biggest in this domain (De Jong, Tillema & Wolfgram, 1985;

Stokking & Leenders, 1988). The central constructs were defined as follows: "Individualizing": allowing for the individual possibilities and promoting the own development of every child; "Differentiation": allowing for individual differences amongst the pupils of a school class with regard to one or several aspects (such as goals, content, learning time, didactic method); "Extra care": spotting and diagnosing of learning problems and trying to remediate these". About 35 predictor variables were selected for the whole project, distributed over the three levels mentioned. Research on IDE in the Netherlands in the 80s showed that relevant predictors are to be found on school level (school size, denomination, school leadership, inservice training and external support) as well as class/teacher level (class size, textbooks, attitudes, inservice training and external support) (Hofman, Van Aalderen & Span, 1984; Span, Abbring & Meyer, 1985; Reezigt, Van Dijk & Bosveld, 1986). These empirical generalizations can be supported by relevant theories on "School effectiveness" (see Van de Grift 1987) and "School development and knowledge utilization" (see Stokking 1988) and can be supplemented by theories on "Diffusion and utilization of information" (relevant predictors: information activity, and location (geographical); House 1974; Rowley & Turner 1976; Stokking & Juli 1986; Stokking & Leenders 1992).

Research methods

The sampling was aimed at achieving a nested structure and certain minimum numbers per nest. For reasons of expected relevance and unambiguous demarcation the division of the Netherlands into the 55 regions of the Regional School Support Agencies (OBDs) out of a ten possible regional divisions was adopted as a starting point. Multi-stage random sampling was applied to 20 OBDs and 1000 primary schools with at least 100 pupils.

Data gathering at the levels of school and class(teacher) was effected by means of two written surveys provided by primary school principals and teachers (with a response of 67% and 61% respectively). The participating schools are highly representative in terms of variables such as school size, denomination and the textbooks in use in reading and arithmetic. The data on the regional level were obtained from already existing research reports.

The data analysis consisted of three parts. First, the psychometric quality of the scales was checked via measurements of internal consistency (Cronbach's alpha). Secondly, the teachers, classes, schools and regions were described in terms of the selected variables (mean and standard deviation). Lastly, multilevel analyses were carried out, by means of the programme ML3 (Prosser, Rasbash & Goldstein, 1991). After removal of 33 schools with less then four teachers and one OBD with less than five schools, 19 OBDs, 215 schools and 1424 teachers remained for the analysis. These numbers were adequate for the intended multilevel analysis (neither too small (Basssiri 1988), nor too large (Snijders & Bosker 1989)).

For the multilevel analyses itself a further selection was made which involved statistically reliable, non-skewed, non-collinear and, with respect to content, most relevant variables. This produced the following *criterion variables*: the degree of individualization and differentiation on the level of class for the subjects of reading and arithmetic, the degree to which teachers allow pupils to work independently and the attitude of the teacher towards IDE. The following *predictor variables* were selected *on the class/teacher level*: the degree to which the teacher actively uses

new information, the frequency with which the teacher gets external (OBD-)support in the implementation of IDE, the number of pupils per class and the relevant year-group(s). *On school level* the following variables were selected: commitment by the school team to extra care in reading, the frequency with which relevant teachers are supported by an OBD in implementing IDE, the degree of educational leadership shown by the school principal and the number of pupils at the school. *On the regional level* the following variables were selected: commitment within the OBD to preventive pupil care, the number of supporters/consultants and the number of schools within the jurisdiction of the OBD, and two demographic indices: degree of urbanization and agglomeration index.
A variance component analysis was carried out to provide an answer to the first research question. In view of the second question predictors were used, on various levels. Finally, random coefficient models were analyzed, with random slopes.

Results
Of the 15 intended Likert scales, 12 scales proved to have a relibility which was satisfactory (a Cronbach's a of at least .70; in fact the a's were between .73 and .91). Some 65-70% of the variance in the variables defined on the class/teacher level occurs on that level, whereas 25-30% (always significant) occurs on the school level and only 5% on the OBD level (significant only for the variable 'independent work').
The degrees to which teachers individualize and differentiate in reading and arithmetic and use independent work correlate positively with the degree to which teachers are active in using new information, the degree of external support of teachers in this area, the number of consultants available to the OBD and the degree to which the OBD is committed to preventive pupil care. Teachers for jounger children and merged-grade teachers individualize and differentiate more often in the teaching of reading. Teachers in the higher grades or in larger classes more often allow pupils to work independently. Frequency of external support for the school team, size of school and demographic variables turn out not to be serious predictors. In a number of cases the slopes differ significantly, in the statistical sense, between schools, but these differences are fairly insignificant in terms of absolute value.

Conclusion and discussion
In terms of individualization, differentiation and extra care, there are fairly large differences in primary education between classes (teachers) as well as between schools, but hardly any between OBD areas. The differences between classes (teachers) may be partly explained in terms of variables which have to do with what teachers, school principals and supporters *actually do*, rather than in terms of various *conditions* such as classroom composition or classroom size, size of school or location of school. The relevance of the variables 'school-size' and 'classroom-size' is not confirmed by this study. Relatively speaking, the educational leadership of school principals proves to be the best predictor, followed by external support for individual teachers (in contrast with external support for teams). In addition, 'information activity of the teacher' proves to have some relevance. These outcomes show that different theoretical approaches,

"School effectivenes" as well as "Information utilization" can contribute insights into what constitutes the relevant factors. The regional differences reported in previous research into diffusion were not confirmed. These differences have obviously disappeared by now.
The variables in the analyses were not centered. Additional analyses with centered variables chiefly changed only the ratio of the remaining unexplained variance on school level to the unexplained variance on class/teacher level.

References

Bassiri, D. (1988). Large and small sample properties of maximum likelihood estimates for the hierarchical linear model. New Orleans: AERA.

Grift, W. van de. (1987). Implementatie van vernieuwingen: de rol van de schoolleider. Den Haag: Vuga.

Hofman, R.H., Van Aalderen, R.J. & Span, B. (1984). Achterblijvers in het innovatie-proces basisschool. Deel II: Omvang, kenmerken en motieven. Haren: RION.

House, E.R. (1974). The politics of educational innovation. Berkely: McCutchan.

Jong, R. de, Tillema, H.H., & Wolfgram, H.P. (1985). Professionele bekwaamheden en nascholingsbehoeften van leerkrachten basisonderwijs. Groningen: RION.

Prosser, R., Rasbash, J. & Goldstein, H. (1991). ML3 Software for three-level analysis. User's guide for V2.0. London: Institute of Education, University of London.

Reezigt, G.J., Van Dijk, M.H., & Bosveld, J.J.F. (1986). Differentiatie op de basisschool. Groningen: RION.

Rowley, J., & Turner, C. (1978). The dissemination of information. London: Deutsch.

Snijders, T.A.B., & Bosker, R.J. (1989). Standard errors and sample sizes for two-level research. Groningen: Sociologisch Instituut.

Span, B., Abbring, I.M., & Meyer, C.J.W. (1985). De opvang van leerlingen met problemen in het regulier onderwijs. Den Haag: SVO.

Stokking, K.M. (1988). National educational policy and external support systems as conditions for curriculum implementation. New Orleans: AERA.

Stokking, K.M. (1992). Wijkende condities. Differentiatie en zorgverbreding in het basisonderwijs, onderzoek en beleid 1984-1990. Utrecht: ISOR.

Stokking, K.M., & Juli, A. (1986). Het functioneren van schoolnabije informatie in het basisonderwijs. Utrecht: ISOR.

Stokking, K.M., & Leenders, F.J. (1988). Effectiviteit van nascholing en nascholingsbeleid. Den Haag: SVO.

Stokking, K.M., & Leenders, F.J. (1992). Heeft verspreiding van informatie zin? Verslag van een onderzoek in het onderwijs. Utrecht: ISOR.

BEST COPY AVAILABLE

BUILDING A COLLABORATIVE SCHOOL CULTURE: A CASE STUDY OF ONE WOMAN IN THE PRINCIPALSHIP

Rouleen Wignall, Ontario Institute for Studies in Education, Canada

The structure of a secondary school typically places the principal at the top of a hierarchy, supported in turn by vice principals and heads of departments. Implicit in this hierarchical structure is the notion of legitimate authority vested in the principal. While a principal may accept responsibility for clarifying the school's mission and for developing a sense of shared vision among teachers, this activity is typically constrained by existing organizational structures built upon subject divisions and their interrelationships.

School subjects bestow identity on many secondary school teachers and give meaning to their work. Subject communities and departments come to comprise work cultures in their own right. [1] In most secondary schools, subject communities and departments also constitute political power bases from which teachers contest issues such as territory, resource allocation, status, and career. This combination of subjects as sources of identity and subjects as sources of power tends to balkanize secondary schools. This balkanization along subject lines, coupled with the individualistic culture of teaching (Lortie, 1975) and the generally accepted hierarchy of authority tends to impede administrative efforts to foster collaboration. Such attempts frequently result in a culture of contrived collegiality (Hargreaves & Wignall, 1989) rather than collaboration.

Most studies of teacher work culture have involved elementary schools which provide relatively small, homogeneous settings. Few such studies have been conducted at the secondary level. Even fewer focus on attempts to develop collaborative work cultures. This paper discusses characteristic features of secondary school work cultures, and the particular features of one secondary school in which the principal successfully encouraged collaboration. Set within a broader study of schools which are typically characterized by balkanized subcultures, this analysis offers insights into the building of collaboration.

This paper examines the relationships among secondary school work culture, structure, and collaboration. Specifically, the paper emphasizes the results of one case study from a larger, qualitative, multi-site research study [2] conducted in Canadian secondary schools reflecting a broad range of school types and communities. [3] The proposed paper, like the broader project, analyzes the reciprocal interrelationships among culture, structure, and change.

The particular case discussed in this paper involves a relatively new secondary school. In the course of the larger study it became apparent that this school

demonstrated the norm of collaboration. Throughout the school's three year existence, the principal has attended to a range of components which appear to affect culture. By respecting the perspectives of other stakeholders, and by addressing issues such as physical plant, allocation of office space, hiring and promotion, and shared decision making the principal has reduced the boundaries which so often exist between departments and has created a collaborative environment in which teachers feel empowered. Teachers at this school appear to be confident in their abilities to face challenges They appear typically less threatened by the prospect of change than teachers working in less collaborative cultures.

Evidence from this case will also be discussed in light of emerging theories of women in positions of leadership. In addition, the paper will indicate how qualitative methods of the kind used in this project allowed the research team to capture the complexities associated with the dynamic evolution of organizational culture. The flexibility inherent in our approach encompassed a range of practices which may not have been anticipated in advance. Furthermore, the face-to-face interaction fostered trust and security, important prerequisites to open frank discussion of issues that can, at times, be personal and sensitive. We have attempted to approach the concept of organizational culture clarifying its meaning for the participants and, exploring its subtleties. Commonly used quantitative methods which seek general beliefs and attitudes typically lead to over-homogenization and exaggeration of the common elements and their importance for participants. Moreover these methods typically preclude extended exploration of themes which emerge during the inquiry which were not anticipated in advance.

Endnotes

- Attitudes to and assumptions concerning the nature of knowledge, teaching and learning, approaches to pedagogy, discipline and so forth tend to vary along subject lines.
- In each school in the larger study, the principal, a vice-principal and at least 12 teachers were interviewed on two occasions. The semi-structured schedule included school organization and decision making practices; teachers' working relations with colleagues; and teacher perspectives and practices concerning pedagogy, subject and career. Teachers were selected according to a range of criteria including status of their teaching subject. Interviews were tape-recorded, transcribed and reviewed for emergent themes and issues.
- Schools were selected to include schools of varied socio-economic status, schools with modified structures including self-paced, individualized, resource-based curriculum, a small school and a school with a planned bias towards subjects which do not have high status. Some schools had pilot projects which had been funded by the Ministry of Education. Others were selected because they had specifically attempted to develop collaborative work cultures among their staffs. The school which provides the focus for this paper was selected because it met 2 criteria - socio-economic status and gender of the principal. It was not selected on the basis of characteristics of work culture.

SCHOOL PRINCIPALS AND THE ENTREPRENEURIAL ORGANIZATION

Hugh Jenkins, & David Evans, Anglia Business School, Cambridge, United Kingdom

Introduction

There is substantial evidence that managers outside the school system are making far reaching changes in the structure, processes and management of their organisations. Typical changes are dumping bureaucracy and hierarchies, flattening structures, building self-managing teams with devolved decision-making, empowering staff, creating open networks of information and breaking down segmentialism. hese changes are intended to achieve high quality and customer satisfaction through increased competitiveness and rapid response to environmental turbulence. School Principals in England and Wales are facing an equally turbulent environment in the light of the rapid and extensive changes in the Education System. These changes in the context of education have been brought about by the ideologies of consumerism, managerialism, centralisation and politicisation. To cope effectively with this new context should not Principals too, be adopting radically new approaches to leadership and to the design of school organisations?

Research Methodology

For the purpose of this paper evidence is presented from detailed case studies of six secondary schools which are known to have made significant changes in their management structures and processes. The six schools are situated in England and Wales and cover a range of geographical and social contexts. All the schools have an enrolment of between 800 and 1200 pupils and have between 40 and 60 teaching staff. The paper, therefore, concentrates on changes in organisational structures and managerial approaches designed to enhance effective learning.

Research Findings

The authors' research into secondary schools in England and Wales over the last 3 years (see Jenkins 1991) indicates that many school. Principals are sticking to traditional hierarchical approaches to mai agement and are still floundering in the attempts to "rethink" management and to a'opt new strategies for effectiveness. However two key results emerged from the r esearch. Firstly school Principals are suffering to a high degree tensions about the need to change. These tensions include:

- The conflict between the leader as hero, the dominant figure in the organisation, and the leader as facilitator, supporter, sharer of power;
- The conflict between managerialism with its emphasis on control of performance through stress on productivity, output measurement and appraisal, and the more professional approaches of caring and valuing of staff;
- The conflict between the educational institution as a bureaucracy with traditional forms of hierarchy and status and the need for the manager to create "flatter" organisation with cross functional managers and high levels of collaboration;
- The conflict between pleasing the customer as arbiter of quality through consumer choice and competition and maintaining the integrity of the organisation through skilled professional leadership;
- The conflict between centralisation and the implementation of political ideologies and the empowerment of staff at the point of service delivery;
- The conflict between the manager as entrepreneur and innovator and the manager as cultural symbol of continuity and social stability.

Secondly a small number of school Principals are beginning to make radical changes in the structure and leadership approaches within the school. They are giving the schools a vision, empowering staff, breaking down bureaucracy, radically changing staff functions, and structuring the school in new flexible ways, while at the same time insisting on high quality and customer care through continuous evaluation and measurement of performance.

The schools in the case studies have got rid of traditional bureaucratic `top-down approaches and the evidence indicates that the changes in structures and processes instigated by the Principals have gone a considerable way towards the creation of the flexible responsive organisation which can meet the demands for quality performance. The case studies also reveal that, in the majority of cases, staff in these schools have responded positively to increased responsibility and to the challenges of empowerment. he flattening of the hierarchical pyramid has released the energy and creativity of staff throughout the school. A number of the schools in the sample are seeking ways of ensuring empowerment which combine power-sharing with structural solutions for managing the school. The Principals are letting go of power and building the organisation through self-managing teams which are responsible for scheduling and designing their own work, evaluating their effectiveness and implementing action for improvement. In these schools the Principals are playing new roles. Like managers in a number of commercial organisations they are now acting as consultants, facilitators, counsellors and empowers.

The school Principals in the case studies are also committed to achieving high quality. The message of quality is highlighted at all times and a strong relationship developed between the professionals and consumers as joint guardians of quality. Above all the Principals are innovators who see change as a normal way of life. They are prepared to constantly seek opportunities for change, attempting new ventures and condoning staff shortcomings and failures in attempts to change. Above all they are expert in creating an entrepreneurial climate by removing inappropriate structures and building trust and confidence in the staff to implement change.

Conclusion

There is evidence that some school Principals are evolving new approaches to the design of school organisations which challenge long-standing ways of thinking about schools as organisations - and which reflect wider developments outside education in the ways organisations are designed and run. It is argued that to ensure effectiveness, school Principals will need to re-think schools as organisations and the case studies show how some Principals have resolved the tensions referred to above and have created entrepreneurial organisations which can respond rapidly and effectively to new challenges.

References

Jenkins, H. O. (1991). Getting it Right:: A Handbook for Successful School Leadership. Oxford: Blackwell.

THE ROLE OF THE LOCAL EDUCATION AUTHORITY IN PROMOTING SCHOOL EFFECTIVENESS AND EVALUATION

Kathryn A. Riley, Institute of Local Government Studies, Birmingham University, Birmingham, United Kingdom

Introduction

Local education authorities (leas) in England and Wales are currently struggling to clarify their educational role. Are they market regulators? Are they advocates for the users of the service? How can they provide a vision and a coherent planning framework for education? What is their role in evaluating, monitoring and developing quality? Do leas have a future at all?

The paper will begin by examining the current policy and political framework of local authority education services in England and Wales and the context for school and local authority issues about monitoring, evaluation and inspection. The main purpose of the paper will be to explore the relationship between the conditions most likely to support the development of effective schools and the role of a locally based education service in promoting those conditions. The paper will illuminate some of the current tensions and problems in that relationship.

In undertaking this analysis, the paper will focus on two themes :

The role of the LEA;

Part one of the paper will explore those elements which contribute to the effectiveness of the lea - such as size, organisation and purpose - by drawing on an evaluative study of the administrative and service consequences of the abolition of the Inner London Education Authority, on 1st April 1990, and the impact of the transfer of education services to the London Boroughs. The fieldwork for this study was carried out through questionnaires, semi-structured interviews and was backed up by a wide range of source documentation. The findings from the case-study will be used to analyse the likely consequences of propsels by british government, to transfer education services from large county authorities to smaller unitary district authorities.

Evaluating effectiveness;

The second section of the paper will focus on one of the major administrative tasks of the lea: accounting for quality, to examnine how the norms and practices of leas can contribute to positive outcomes in schools. In particular, the paper will examine the relationship between inspection, school self-evaluation, research and development and education indicators. To support this analysis, this section of the paper will draw on a research project undertaken for the Local Government Management Board on education indicators in 1991 and a series of follow up

workshops in the Spring of 1992, with headteachers, inspectors and senior education officers on quality and inspection in the education service.

Issues and themes

In exploring issues about the size and purposes of leas and their role in accounting for quality, a number af themes and tensions emerge:

- between the benefits offered by a strategic education authority versus those offered by a smaller unitary authority;
- between external validation of schools versus internal self-evaluation and review;
- between the need to demonstrate the value added by the school to the process of teaching and learning versus value added by the lea.

The concluding part of the paper will examine these themes and reflect on the current and future education agenda facing schools in England and Wales over the next few years, examining the question: what type of system will best support the development of quality in schools?

THE EFFECTS OF CATHOLIC SECONDARY SCHOOLING ON PUPIL PERFORMANCE

P. Daly, Nothern Ireland

The paper will provide a comparative account of the two school systems in Nothern Ireland, often described as Catholic and Protestant, in effectiveness terms. The main output measure will be academic performance in public examinations taken at the end of secondary schooling. The paper extends earlier work of the author (Daly, 1991) by focussing on Catholic schools in the light of access to a second and more up-to-date data set and to recent Scottish findings.

It will have six main sections:

- A brief historical review of Catholic school provision since the establishment of the state in 1921.
- Discussion of known raw performance differences between the two systems, ie. unadjusted for pupil-intake differences.
- Presentation of intake-adjusted attainment measures for each system, based on a re-analysis. The Northern Ireland Council for Education Research (NICER) has made two pupil-cohort data sets available for re-analysis. The data describe aspects of family background, achievement at the first-year stage of secondary schooling and attainment in public examinations taken at the end of secondary schooling. This re-analysis of data takes an hierarchical linear modelling approach, using the HLM programme developed at the University of Chicago (Bryk et al., 1986). Software for multi-level modelling was not available to the original NICER researchers.
- These new Northern Ireland results will be compared to recent estimates of Catholic school effects in Scotland, which were also derived from the HLM programme (Willms, 1989).
- Recent statutory changes in regard to the organisation of schooling in Northern Ireland and their implications for the future delivery and monitoring of the education service (schools) will be considered in the light of this study.
- The final section will provide a summary and conclude with policy recommendations.

References

Bryk, A. S., Raudenbush, S. W., Seltzer, M., & Congdon, R. T. (1986). An introduction to HLM: Computer programme and users guide.

Daly, P. (1991). How large are secondary school effects in Nothern Ireland? School Effectiveness and School Improvement. Vol. 2, No. 4.

Willms, J. D. (1989). Pride or prejudice?: Opportunity structure and the effects of Catholic schools in Scotland. Paper presented at a meeting of the Standing Commission on Human Rights in Northern Ireland. Belfast.

LINKING TOGETHER SCHOOL EFFECTIVENESS KNOWLEDGE AND SCHOOL: IMPROVEMENT PRACTICE

D. Hopkins, Cambridge Institute of Education, Cambridge; D. Reynolds, School of Education, University of Wales College of Cardiff, Cardiff, United Kingdom

At present, it is argued there exist few close links between the 'specialties' of school effectiveness and school improvement. With the exception of the United States, where school effectiveness research findings are now being utilised in over half of American school districts, there is a notable absence of the direct transplantation of school effectiveness research into practice. In all societies, including the United States, the two academic communities are very separate, as shown by such factors as citation patterns.

The school improvement paradigm of the 1970s and 1980s is outlined, with its emphasis upon the importance of teacher professional 'lore', ownership, changes in processes, the utility of qualitative methods and the importance of a 'goals debate' as part of the process of school improvement. By contrast, the school effectiveness paradigm was outcome orientated, quantitative in orientation, eschewed any goal discussion and focussed upon organisational and behavioural factors within schools, rather than attitudes or attitudinal change.

The possible contributions of each of the above groups to the needs of the other are outlined. School effectiveness research can, for school improvement:

- Generate fine grained, qualitative case studies of exceptional schools;
- Focus upon organisational processes;
- Generate dynamic, over-time analyses of schools as organisations;
- Assess which process variables are causes of effectiveness and which results of being effective.

School improvement practice can, for school effectiveness research:

- Measure the impact of organisational improvement upon pupil outcomes, to see which school process factors may be causal;
- Bearing in mind the implications of multi-level modelling and its results on within school differentiation, focus and target improvement not at a 'whole school level' but on different organisational practices appropriate to different pupil groups;
- Focus upon the level of instructional processes, to fill this neglected 'black box' within the school effectiveness research enterprises of certain countries (eg Great Britain).

Examples are finally given of three programmes which represent a blurring of distinctions concerning 'effectiveness' or 'improvement', and which exhibit a potential blending together of the two approaches. These are:

- The Halton Board of Education's school development initiatives in Canada;
- The 'Improving the Quality of Education' project at the Cambridge Institute of Education, UK;
- The 'School Change Agents' Programme at the University of Wales, Cardiff, UK.

THE EFFECTIVENESS OF THE IMPLEMENTATION AND THE USE OF A COMPUTER ASSISTED ATTENDANCE REGISTRATION SYSTEM

K.Tj. Bos & A.J. Visscher, Centre for Applied Research on Education (OCTO), University of Twente, Enschede, The Netherlands

Abstract

Effective schools studies have shown relationships between the amount of truancy in schools and their organizational characteristics. Absent pupils reduce their effective instruction time and persistent absenteeism is assumed to be related to drop out. One possible way to combat truancy is the use of a computer-assisted absence registration system (ARS). The implementation and use of an ARS in thirty secondary schools in four big Dutch cities has been evaluated in a longitudinal study with a pretest-posttest-control group design. Important factors for ARS use seemed to be 'motivation' and 'encouragement'. A significant effect of the use of ARS on the truancy rate could not be demonstrated. Nevertheless the majority of the experimental schools reported some other strong positive effects and no serious negative effects of implementing and using ARS.

Introduction

The advantages of the use of the computer in terms of an increase of efficiency and effectiveness have lead to the widespread use of computer technology for administrative and management activities in schools. This paper concerns the evaluation of the implementation and use of a computer assisted 'attendance registration system' (ARS). In the Netherlands, as well as in Great Britain and the U.S.A. truancy and student drop out are serious problems in secondary education. Effective schools research has shown the relationship between the truancy magnitude of schools and their organizational characteristics (e.g. Rutter, 1979; Stoel, 1986). Absent students reduce their effective instruction time and persistent absenteeism is assumed to be related to drop out. Therefore it is important to take the truancy problem seriously.

One way in which schools can combat absenteeism is to register the attendance of students by means of ARS. The main advantage of ARS is that it generates truancy figures and statistics efficiently and timely. As a result of that schools are able to react to truancy more quickly and adequately, which might contribute to a reduction of the truancy rate. Moreover, ARS enables the analysis of truancy trends and patterns as well as relations between truancy and other variables like timetable characteristics, teachers, subjects and test scores. On the basis of this

information schools can develop school anti truancy policies to reduce the magnitude of absenteeism.
With this in mind the Dutch government in 1988 started a project in which thirty secondary schools in four cities participated voluntarily. These schools implemented ARS (hard- and software) in their organisations.
At the same time a three year longitudinal evaluation study was started to examine the effect of ARS use on truancy rates, as well as on other effects. The results of this study will be presented in this paper.

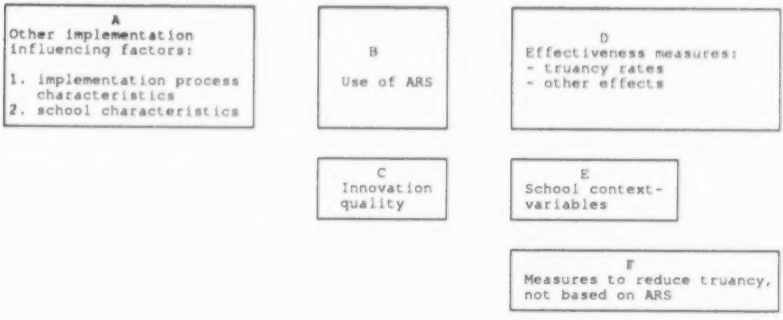
Questions

The central questions of this research project can be stated as follows:

- Which factors stimulate or hinder a successful implementation of ARS?
- To which extent can the truancy magnitude be reduced by means of ARS and which other effects (than changes in truancy rates) has ARS use?

To answer these questions a research framework has been constructed on the basis of literature research. The research framework for the ARS evaluation study is shown in figure 1.

Figure 1: Research framework for the ARS evaluation study



A crucial variable in the study is 'the use of ARS' (block B). This variable can be affected by characteristics of the innovation process (A1): for example by the degree to which schools are coached during the implementation process, the innovation resources schools receive and the degree to which the principal encourages school staff to innovate. School characteristics (A2) can also influence the use of ARS: one can for instance think of the degree of pupil counselling and the degree of teacher support, the innovation motivation of school staff and the resemblance between ARS procedures and the old, manual way of attendance registration of a school. It is assumed that the degree of ARS use is affected by the quality of ARS (the quality of the hard- and software of ARS, block C) as well. The effectiveness of using ARS has been measured by determining changes in truancy rates between 1988 and 1990, 1990 and 1991 and between 1988 and 1991

(block D). It was expected that truancy rates would decrease after ARS was implemented. However, installing and using computers often also goes along with unplanned positive and/or negative effects. For that reason it was also investigated to what degree such effects were perceived by school staff.
In order to be able to attribute a measured effect to ARS use, the influence of school context variables (E) such as school size, percentage of non-Dutch pupils and the social economic background of the pupils (SES) had to be controlled for. The same goes for variable F, the number of truancy reducing measures, taken by schools, that are not based on ARS.

Method
The ARS evaluation study has a longitudinal quasi experimental design with one pretest and two posttests. This design can be marked as a pretest-posttest-control group design. Data have been collected in an experimental group which consists of thirty schools for vocational and general secondary education in four big Dutch cities which are voluntarily participating in the ARS-project. Data were gathered before ARS implementation in April 1988 (the pretest) and in April 1990 (posttest one) and April 1991 (posttest two).
In the literature about truancy a great diversity of truancy definitions, methods to measure them and of ways of data collection can be observed (e.g. Rutter, 1979; Galloway, 1985; Reid, 1985; Grimshaw & Pratt, 1986; De Vries, 1987). As a consequence of that different studies result in truancy figures that are very difficult to compare. In most studies truancy is not measured very exactly. In the ARS evaluation study it was tried to measure truancy with a strict method. The attendance data were gathered during three days in one week in April 1988, April 1990 and April 1991 during all lessons (ca. 8500 in each year), except those given in examination classes. During each lesson, teachers registered names of pupils who did not attend their lessons on special forms. Subsequently school staff, responsible for absence handling, reported which pupils were absent with permission, and during which lesson periods. By comparing both sets of data the percentage of truancy per day, per grade, per subject, per timetable-hour, etcetera could be calculated.
The control group in this study consisted of nineteen schools for secondary education. In these schools, which are of the same size as the experimental schools and which are located in the same cities as the experimental schools, student attendance is registered manually. Comparing changes in truancy rates between 1988, 1990 and 1991 in the control group with those in the experimental group was necessary to avoid false conclusions about the effectiveness of using ARS.
The other variables (features of the school organisation, implementation process, school context characteristics, ARS use, innovation quality, etcetera) were measured in 1988, 1990 and 1991 by interviewing school principals and clerical staff. Besides, questionnaires were handed out to the management, teachers and all other staff members involved in attendance registration (e.g. school caretaker, clerical staff etc).
Data analyses resulted in output statistics, like frequencies and association measures such as Pearsons Product Moment correlations and non parametric correlations.

The effectiveness of the use of ARS was analysed by means of multiple regression analysis and analysis of variance. In these analyses dependent variables were variables either from block B or block D (figure 1). The predictive variables for Block B (ARS use) were the variables from block A and block C. In the analysis of the predictive value of ARS use (block B) for changes in truancy rates (block D), the context variables from block E and F should have served as control variables.

Results

Regression analysis on the 1990 data showed that among others the motivation towards ARS use and the degree to which schools meet the requirements for ARS use are important factors that influence ARS use. Examples of such requirements are the degree to which schools consider the school characteristics as one of the truancy causes and the degree of pupil guidance. These variables explain 48 per cent of the variance in ARS use. The same analysis on the 1991 data did not result in any predictors of ARS use ($p < .05$).

Regarding truancy rate differences (between 1988-1990, 1990-1991 and between 1988-1991) in the experimental and control group, on the basis of variance analysis it can be concluded that although these differences are stronger in the experimental schools than in the control group, a significant effect ($p < .05$) of the use of ARS on the truancy rate could not be demonstrated. An explanation for this might be that most schools only used ARS in a clerical, registrational way and did not develop anti truancy policies on the basis of ARS information. Results indicate that some context variables like 'ethnic background of puplis' (percentage non-Dutch pupils in school) and school size are important for the degree of truancy and the degree to which truancy can be reduced by means of ARS use.

Moreover, the majority of ARS schools reported some other strong positive effects of using ARS: schools have a better insight in truancy trends and patterns and in relations between truancy and other variables (for example timetable characteristics and testscores), the quality of absence registration and handling has improved and computing truants for pupil reports requires less work than when ARS was not used. On the other hand schools did not mention any serious negative effects of using ARS.

References

Galloway, D. (1985). School and persistent absentees. Oxford: Pergamon Press.

Grimshaw, R. & Pratt J. (1986). Counting the absent scholars: some implications for managerial practice arising from a survey of absenteeism in a city's secundary schools. School Organization, 6, 1, 155-173.

Reid, K. (1985). Truancy and School Absenteeism. London: Hodden and Stoughton.

Rutter, M., Maugham B., Mortimore, P., & J. Ouston, J. (1979) Fifteen thousand hours. Secondary school and their effects on children. London: Open Books.

Stoel, W.G.R. (1986). Schoolkenmerken en het gedrag van leerlingen en docenten in het voortgezet onderwijs. Groningen: RION.

Vries, G.C. de (1987). Verzuim, uitval en de bijdrage van de school. [Truancy, dropout and the contribution of the school] Pedagogisch Tijdschrift, 12, 2, 94-104.

BUILDING A COLLABORATIVE SCHOOL CULTURE: A CASE STUDY OF ONE WOMAN IN THE PRINCIPALSHIP

Rouleen Wignall, Ontario Institute for Studies in Education, Canada

The structure of a secondary school typically places the principal at the top of a hierarchy, supported in turn by vice principals and heads of departments. Implicit in this hierarchical structure is the notion of legitimate authority vested in the principal. While a principal may accept responsibility for clarifying the school's mission and for developing a sense of shared vision among teachers, this activity is typically constrained by existing organizational structures built upon subject divisions and their interrelationships.

School subjects bestow identity on many secondary school teachers and give meaning to their work. Subject communities and departments come to comprise work cultures in their own right. [1] In most secondary schools, subject communities and departments also constitute political power bases from which teachers contest issues such as territory, resource allocation, status, and career. This combination of subjects as sources of identity and subjects as sources of power tends to balkanize secondary schools. This balkanization along subject lines, coupled with the individualistic culture of teaching (Lortie, 1975) and the generally accepted hierarchy of authority tends to impede administrative efforts to foster collaboration. Such attempts frequently result in a culture of contrived collegiality (Hargreaves & Wignall, 1989) rather than collaboration.

Most studies of teacher work culture have involved elementary schools which provide relatively small, homogeneous settings. Few such studies have been conducted at the secondary level. Even fewer focus on attempts to develop collaborative work cultures. This paper discusses characteristic features of secondary school work cultures, and the particular features of one secondary school in which the principal successfully encouraged collaboration. Set within a broader study of schools which are typically characterized by balkanized subcultures, this analysis offers insights into the building of collaboration.

This paper examines the relationships among secondary school work culture, structure, and collaboration. Specifically, the paper emphasizes the results of one case study from a larger, qualitative, multi-site research study [2] conducted in Canadian secondary schools reflecting a broad range of school types and communities. [3] The proposed paper, like the broader project, analyzes the reciprocal interrelationships among culture, structure, and change.

The particular case discussed in this paper involves a relatively new secondary school. In the course of the larger study it became apparent that this school demonstrated the norm of collaboration. Throughout the school's three year existence, the principal has attended to a range of components which appear to affect culture. By respecting the perspectives of other stakeholders, and by addressing issues such as physical plant, allocation of office space, hiring and promotion, and shared decision making the principal has reduced the boundaries which so often exist between departments and has created a collaborative environment in which teachers feel empowered. Teachers at this school appear to be confident in their abilities to face challenges They appear typically less

threatened by the prospect of change than teachers working in less collaborative cultures.

Evidence from this case will also be discussed in light of emerging theories of women in positions of leadership. In addition, the paper will indicate how qualitative methods of the kind used in this project allowed the research team to capture the complexities associated with the dynamic evolution of organizational culture. The flexibility inherent in our approach encompassed a range of practices which may not have been anticipated in advance. Furthermore, the face-to-face interaction fostered trust and security, important prerequisites to open frank discussion of issues that can, at times, be personal and sensitive. We have attempted to approach the concept of organizational culture clarifying its meaning for the participants and, exploring its subtleties. Commonly used quantitative methods which seek general beliefs and attitudes typically lead to over-homogenization and exaggeration of the common elements and their importance for participants. Moreover these methods typically preclude extended exploration of themes which emerge during the inquiry which were not anticipated in advance.

Endnotes

- Attitudes to and assumptions concerning the nature of knowledge, teaching and learning, approaches to pedagogy, discipline and so forth tend to vary along subject lines.
- In each school in the larger study, the principal, a vice-principal and at least 12 teachers were interviewed on two occasions. The semi-structured schedule included school organization and decision making practices; teachers' working relations with colleagues; and teacher perspectives and practices concerning pedagogy, subject and career. Teachers were selected according to a range of criteria including status of their teaching subject. Interviews were tape-recorded, transcribed and reviewed for emergent themes and issues.
- Schools were selected to include schools of varied socio-economic status, schools with modified structures including self-paced, individualized, resource-based curriculum, a small school and a school with a planned bias towards subjects which do not have high status. Some schools had pilot projects which had been funded by the Ministry of Education. Others were selected because they had specifically attempted to develop collaborative work cultures among their staffs. The school which provides the focus for this paper was selected because it met 2 criteria - socio-economic status and gender of the principal. It was not selected on the basis of characteristics of work culture.

SCHOOL IMPROVEMENT THROUGH SYSTEMATIC FEEDBACK OF PUPIL LEVEL DATA AT THE SCHOOL AND CLASSROOM LEVEL

H.P. Brandsma, Centre for Applied Research on Education (OCTO), University of Twente, Enschede; P. Edelenbos, Institute for Educational Research (RION), University of Groningen, Groningen, The Netherlands

Introduction

Although it might be true that school effectiveness research has become a major tradition in educational research it has been stated moreoften that these studies suffer from serious theoretical and methodological flaws. Summarizing these criticisms one can say that most of the research sofar can be characterized as a fishing expedition with large drawing nets in which the explanatory variables are sometimes caught and sometimes not. The major criticism is that no explicit causal theoretical model has been suggested in which all of the explanatory variables can be ordered in a theoretically and empirically logical way. The methodology has been inductive and emperistic to a large extent. Due to the correlational approach of most studies no causal interpretations could be made. This line of research however has made it possible to formulate more specific hypotheses that can be put to an empirical test. For this reason it is suggested that the school effectiveness research should follow a deductive line of inquiry in which some of the formulated hypotheses can be studied in depth by making use of more sophisticated designs.

Taking these criticisms seriously an experiment was started in which some of the promising effective school characteristics are implemented in the school organization, by means of a training program. The implementation process is geared at two distinctive levels within the school; the level of the school principal and the level of the teacher.

Evaluative potential and feedback.

It has been stated moreoften that one of the most promising factors for improving the effectiveness of a school is the 'evaluative potential' of the schoolorganization. Using the outcomes of evaluative procedures at different levels in the school organization can be regarded as the steering mechanism for enhancing the quality of the educational process and the educational outcomes. Evaluative procedures can be used in several ways with different goals in mind. In our view the evaluative procedures should direct itself primarily towards the educational outcomes i.e. the learning of students.

Knowledge about the strong and weak points in the quality of the learning outcomes is a very powerful tool for specific schools in determining there schoolgoals and where they should direct their resources at.

A necessary condition for this however is that schools need to be equiped with an evaluative instrument that makes it possible to detect their strong and weak points as compared to their internally stated goals (internal quality), but also as compared to their relative position to other schools (external quality). With regard to the

determining of the external quality the evaluative instrument needs to be constructed is such a way that a 'fair' comparison between schools is made possible. A fair comparison between schools is made possible if there is a control for the influence of the background characteristics of their student population in terms of socio-economic status, ethnicity and prior ability, on the learning outcomes. This concept of 'value added quality' of schools is the baseline and starting point of the experiment that is aimed at improving the quality of schools, i.e. enhancing the effectiveness of the school.
Knowledge about one's quality however is only the starting point for the school improvement process. It provides the school with an incentive and gives guidelines as to where the school needs improving but doesn't say how to do it. Notwithstanding the fact that the learning process primarily takes place at the classroom level the school improvement process can be geared at two levels within the school; the administrative level (i.e. the principal) or the instructional level (i.e. the classroom teacher). This experiment is aimed at these two distinct levels separately. For both levels a different training program is developed.

Principal training
At the school level the central role of the principal as an educational manager and internal staff developer is stressed. In the training program the principal is fed with the outcomes of the evaluative procedure that gives him insight into the relative position of the school at large, the different teachers and the individual students. The principal is taught how to interpret the quality indicators and how to give feedback to the individual teachers with the aim of improving their instructional procedures. In this sense the principal is an intermediary factor in the school improvement process.
The training program for the principals consists of four separate days spread over two school years and are planned in such a way that in every session the principal has insight into the actual outcomes of their students at that moment.

Teacher training
The teacher training courses for mathematics teachers have two main goals: a. teaching the teachers to structure learning tasks and b. teaching the teachers to provide meaningful feedback. The teacher training programme fits the framework of a new mathematics curriculum (basic education for secondary education) with new educational objectives and instructional demands. The organization of the teacher training course, which aims at effective instruction, is based on five important instructional features:

- monitoring and revision of subject matter of previous lessons as an introduction to new subject matter, by means of "organizers";
- gradual presentation of new subject matter, while taking frequent problems into account;
- supervision of pupils during practice;
- providing feedback and corrections;
- assessment of achievements for each subject matter unit.

During the first year of the teacher training course (1991-1992) attempts are made to alter or improve three skills of the teachers:

- structuring learning tasks;

- providing pupils with systematic and meaningful feedback;
- acquiring a teaching model (or variant models) which encourages the active learning time of pupils.

Management and content of the teacher training meetings serve as a model for mathematics teaching that is more effective. In itself this is not an aim, however, it is a prior condition with respect to planning and realisation.
Classroom consultation takes place each time two training sessions have been completed by the participant in the training programme, in support of the training sessions and for the advancement of reflection upon personal educational practice. For that purpose a lesson of the first secondary school year is visited by a classroom consultant and discussed afterwards from the perspective of the previous training sessions. Hence, feedback is provided to the participants.

The experiment
The improvement program described above can also be seen as an experiment between two lines of theory within the effectiveness model i.e. school effectiveness as compared to teacher effectiveness. The aim of the experiment therefore is not only to determine the effectiveness of the programs at the school and teacher level as compared to a control group of schools but also whether it is more effective (and cost efficient) to direct the improvement process at the principal or the teacher.
The program is directed at improving the student learning in mathematics at schools for Lower General Secondary Education (MAVO). The students in the experimental groups are tested five times over a two year period in their mathematics proficiency. The students in the control group of schools are tested twice (pre test at the start of schoolyear 1 and a post test in schoolyear 3). These test scores for the experimental groups are used primarily for providing data feedback to the principals and mathematic teachers who participate in the training program. These data will also be used in evaluating the effects of both improvement programs as compared to the scores in the control group.

SCHOOL IMPROVEMENT: THE AGENDA FOR THE 1990'S

Robert Bollen, Algemeen Pedagogisch Studiecentrum, Amsterdam, The Netherlands

In the 1980s School Improvement has been put on the agenda by the OECD (Organisation for Economic Cooperation and Development) by means of the International School Improvement Project (ISIP). This project produced a series of 14 books and technical reports of which "Making School Improvement Work" (1985) gives the conceptual framework that was used in the other publications. Six years after ISIP came to an end the world is dramatically changed and it is interesting to see whether the basic message of the book and the project is still applicable to the new circumstances.
In the 1990s educational budgets are indeed as tight as foreseen in the decade before, which explains the great interest of policy makers in the possibilities of the effective school. As political favoured change is not likely to happen without any

external pressure, budget cuts combined with a greater autonomy of schools seem to be the political key to improvement. Concepts as site–based management, teacher empowerment and restructuring schools seem to be closely related to the political demand for more effectiveness. The problem is not so much to have a vision or even a sharp picture of the effective school determined by a varying number of variables, but how to transform the ineffective school into the effective school.
Looking back in the opus magnum of ISIP it still is as clear as it was, that improving schools is a multi level process that finally must be implemented and institutionalized within individual schools and classrooms. It also is clear that though schools are the unit of change, schools have strong relationships with their environment, though in different ways at the different levels of education. And finally there is strong evidence that external support in various forms is an indispensable ingredient for real improvement.
The increase in autonomy as is apparent in the current political situation in many countries has made a strong difference in the unit of change. It shifts the role of the principal from a potential instructional leader to a more management oriented leadership. The need to be more productive if not effective will certainly influence teacher behaviour, but not always in a positive way.
It also makes a difference with regards to the catching area and the selective dimensions of the school. And finally it sharpens as well the other side of the picture of autonomy: the need to be accountable. That means that the relationships with the environment may change in a very fundamental way as already can be observed in many countries.
External support itself is also involved in a process of change,driven by the same political forces that so strongly influences their potential clients: the schools. In countries with a well elaborated support system like in the Netherlands, meant to serve both the central authorities and the schools, implementing educational policies by processes of adoption and adaption at the school level, the consequences of a market oriented support system become more and more visible and need to be thought over.
In the School Improvement movement there has been a strong emphasis on the process which is also visible in its methodology: using the case study as a vehicle to improve the knowledge about improvement and producing conceptual frameworks as a base for strategies for improvement. Looking at the methodological aspects of School Improvement knowledge it is obvious that using the case study as vehicle implicates that the knowledge base cannot be stable. After several years case studies become more and more outdated and finally turn into fairy tales, though that takes quite a long time. Of course it is true that quantitative data also loose their reliability, but that often is not so obvious as quantitative data are of a far more abstract nature than case studies or qualitative data. Strategies and conceptual frameworks, derived from the school improvement knowledge body, still may look very reliable because of their abstract nature, but like quantitative data they may be only applicable to the world of the past.

Thinking about an agenda for the nineties for School Improvement research and activities, it is clear that current strategies and frameworks need to be tested again and again. But there is an important and relevant distinction between the starting

position of ISIP and of what is to be undertaken in the nineties. ISIP went to the schools to learn about what was going on and to learn from success and failure in School Improvement with the explicit intention to produce more or less coherent theories, frameworks and strategies to be used by politicians, authorities, support systems and by practitioners in the educational field. A phase of learning by doing needed a phase of reflection to shift implicit understanding of School Improvement up to the level of explicit understanding and to the even higher level of understanding the meaning of change itself. By now everybody engaged in research and policy making declares that we know a lot more about change than a couple of decades ago. So the next thing to be done is to see whether those statements indeed are true or false. Instead of looking at change to produce School Improvement we now need to look at School Improvement, planned and systematic sustained efforts aiming at change in learning conditions and other related internal conditions in one or more schools, with the ultimate aim of accomplishing educational goals more effectively. This definition of School Improvement as formulated in the ISIP book "Making School Improvement Work" needs to be tested itself as well as the frameworks and strategies that emanated from it. If educational goals need to be accomplished more effectively and that is what is done in "effective schools", we need to know how to do that by improving the schools. A careful look at cases in which this was intended and succeeded or failed and a close co–operation of researchers and change facilitators must be on the agenda of the School Improvement movement with the ultimate aim to convince the policy makers whowith no doubt will have to take responsibility for the quality of education in the next century.

INTERNATIONAL STUDIES ON SCHOOL EFFECTIVENESS

B.P.M. Creemers, Rion, Institute for Educational Research, University of Groningen, Groningen, The Netherlands

School effectiveness research in the USA resulted in the so-called factor models for school effectiveness. The most famous one was the so-called five-factor model by Ron Edmonds, but there are others with seven, eight and eleven factors. Research in other countries could not provide evidence for all the factors, partly because of methodological problems and the conceptual framework and partly because of contextual differences between the different countries. In this symposium we will look at the results of some international studies in which information is gathered about factors that can contribute to school effectiveness. The problem of the earlier studies like IEA is that they are not specifically directed towards questions about school effectiveness, so in comparative research proxy variables for school effectiveness have to be developed. The results of two of these secondary analyses of IEA data will be presented. Studies like the OECD indicator study and the International School Effectiveness Research Program are in a stage of development. A progress report of the planning will be presented. Furthermore, a study comparing the differences between educational systems which can explain the results of education will be discussed.

EFFECTS OF SCHOOL SIZE IN SWEDEN, THE NETHERLANDS AND THE USA

Hans Luyten, Department of Education, University of Twente, Enschede, the Netherlands

Presently school size is a topic of governmental concern in the Netherlands. This is mainly because it is widely believed that the creation of larger schools offers possibilities for cost reduction. In this paper the results of an investigation into the relation between the size of a school and the achievements of its pupils are reported from an international perspective. The datafiles which were used relate to Dutch, Swedish and American pupils. The data originate from two international studies: The Second International Mathematics Study (SIMS) and the Second International Science Study (SISS). The SIMS-files which were analysed contain Dutch, Swedish and American data. The SISS-files contained only Dutch data. The SIMS-data were collected in 1980 (Sweden), 1981 (Netherlands and the USA) and 1982 (USA). The mean age of the pupils in these surveys is about 14 years. The

SISS-data were collected in 1984 and relate to pupils with a mean age of 15 years and six months. In SIMS achievement was measured by means of a multiple choice mathematics test. The multiple choice test used in SISS consists of items about physics, chemistry, biology and earth science.
In the analyses the following variables served as covariates:

- Sex;
- Social economic status;
- Achievement motivation;
- Cognitive aptitude;
- Type of education;

It was also investigated if interaction-effects could be discerned between school size and one of these first four covariates. If interaction-effect would be revealed this would imply that the effects of school size on student achievement differ for certain groups of pupils. In the analyses school size was operationalised as a categorical variable. Thus it would also be possible to reveal non-linear relations between school size and pupil achievement.
None of the analyses revealed a statistically significant effect of school size on pupil achievement. Of the several dozens of interaction-terms which were examined only one showed a statistically significant effect on pupil achievement: In the Netherlands the girls in schools with at least 360 but less than 500 pupils got better results on the mathematics test than there male classmates.

STANDARDIZED ACHIEVEMENT TESTING IN DIFFERENT CURRICULAR SETTINGS

W.J. Pelgrum, Department of Education, University of Twente, Enschede, The Netherlands; D.M. de Haan, Open University, OTIC, Heerlen, The Netherlands

Introduction
Although the past decade has shown in many countries an increased interest in monitoring educational progress by comparing achievement measures over time and between nations, methodological sophistication has lagged behind to warrant conclusions commonly based on these comparisons. Especially the interpretation of scores on standardized achievement tests in terms of educational effectiveness does not take into account curricular variation that may lead to differences in overlap of the test and curriculum for groups of students which are compared. This paper examines empirical evidence related to this problem, the nature and validity of measures for registering test-curriculum overlap, and it discusses results from a study for improving these measures.

Analyses of "old" measures
It has been argued that a number of basic conditions are essential for adequate educational assessment. Educational assessment is almost by definition a large-scale enterprise, involving large samples of schools, teachers and students. Given

the fact that many actors are involved in realizing and appraising the output of an educational system, from a communicative point of view it is essential that outputs of the main sub-systems are registered. A minimum requirement would be the measurement of the intended, implemented and realized curriculum as indicators of the output at macro-, meso- and micro-level of the system. An important problem to be solved is the adequate measurement of these different outputs. In many large-scale studies conducted by International Association for the Evaluation of Educational Achievement (IEA) a choice is made for item-based measurement of the different outputs as a basis for processing and evaluation of the measures of the realized curriculum. Pelgrum (1990) addresses the question of how valid and how applicable item-based measures of the implemented curriculum are (as used in the IEA studies). In order to shed light on the validity issue, secondary analyses of data collected in the Second International Mathematics Study (SIMS) and the Second International Science Study (SISS), conducted by IEA, were undertaken. The implementation measures collected in SIMS consisted of teacher and student ratings of test items on the question of whether the corresponding subject matter was taught, whereas for the Netherlands some additional measures of the same variable were collected. In the Dutch part of SISS the same teacher ratings as in SIMS were used. In addition, data for one country in SISS were available in which we use a multidimensional item-based rating of implementation by teachers. The results presented showed that in general the validity of the ratings by teachers is promising. The ratings of teachers corresponded with the content of the textbooks they used, while factor analyses of the ratings reproduced the (curriculum) structure of the Dutch school system. The differences between Dutch school types in subjects included in the lesson tables was reflected in the ratings. Comprehensive school systems showed much less differences in implementation than non comprehensive systems. However, it was also noted that teachers tended to under-rate the amount of subject matter presented to students. Furthermore, it was shown that for particular subsets of items teachers in the Netherlands made serious mistakes in Judging whether the corresponding subject matter had been taught before the date of testing. Pelgrum (1990) concluded that continuing work is needed to study the feasibility of item-based measures of the implemented and intended curriculum.

Analyses of "new" measures

De Haan (1992) addressed the question how ~he existing IEA measures might be improved. She constructed a revised version of the item-based measures as used in the IEA studies by not only asking teachers to judge for each item the difficulty and whether the corresponding subjectmatter was taught. but she also added ratings regarding the suitability of the terminology and the format of the item, such as terminology, format or symbols. She called this instrument D-TCO (Detailed Test Curriculum Overlap). Furthermore, she used a so called H-TCO (Holistic Test Curriculum Overlap) instrument asking teachers to select items they judged fair to administer to their students. In a pre- post test design teacher, textbook, and student ratings were collected. The results of this study are summarized below. With regard to the *reliability of* TCOjudgements collected with the D-TCO instrument it can be concluded that the stability of these ratings is acceptable if the Judgements of teachers are recoded to a dichotomy of whether or not an Item is

taught before the date of testing (D-TCO Judgements). With the D-TCO instrument, 95% of the items that were judged as taught at the pretest, were also judged as taught at the posttest. With the H-TCO instrument this percentage was 92%. Hence the stability of the H-TCO Judgements seems also acceptable. ~he stability of the unrecoded D-TCO judgements (that is, the judgements of whether an item is taught In a specific time period) is considerably lower. Judgements of whether an item was taught in elementary school are found to have the lowest reliability. With regard to the *construct validity of* the D-TCO instrument as well as of the H-TCO instrument, it can be concluded that on an aggregated level (for groups of teachers and items), the construct validity Is reasonable. At the pretest as well as at the posttest, the D-TCO judgements and the H-TCO Judgements correlate in general the highest with each other (varying between r-.78 and r=.91) and correlate lower, but still significantly, with the textbook based TCO judgements (varying between r=.50 and r=.76) and with the student based TCO judgements (varying between r=.56 and r=.84). A comparison of the absolute differences between percentages of TCO based on different approaches, showed that the percentage of DTCO judgements is significantly different from the textbook based TCO judgements. It was supposed that by adapting the results of a textbook analysis for each individual teacher to differences in textbook use (as measured with the teacher questionnaire and with the registration forms filled in by teachers during the period between pretest and posttest), the textbook approach could be an appropriate measure of the operational curriculum. A possible explanation for the discrepancies between D-TCO and textbook based TCO is that the textbook based approach, although adapted to individual differences~ between teachers, is still predominantly referring to the formal curriculum, while the D-TCO measure reflects the operational curriculum. It was also shown that the different TCO measures were significantly higher at the posttest than at the pretest. Hence it seems to be possible to detect changes in a curriculum over time by use of TCO measures. No significant differences are found between D-TCO-measures of different school types. An explanation can be found in the selection of items: the items that were judged by teachers of both school types did not differentiate enough between the curriculum of both school types. The examination of the construct validity at a more specific level showed that the D-TCO instrument as well as the H-TC~O instrument are less convergent with the other approaches: analysing the convergence of different approaches for each teacher individually showed that DTCO judgements were convergent with H-TCO Judgements and with the textbook based TCO judgements. But the average percentage of convergent judgements per item of both the D-TCO and the H-TCO measure varied between 75% (sd=17) and 84% (sd=ll). The size of the standard deviations are quite high, which means that there is a great variation In convergency over items. This provokes the idea that the validity of Judgements varies over different Items. At this specific level, student-based rating appeared to be less convergent with the other approaches. With regard to the Judgement of whether an item deviates on one or more of the characteristics, It can be concluded that in general D-TCO Judgements are convergent with textbook based TCO judgements. However, if one only looks at the convergence of Judgements of items that are Judged as deviating on a specific characteristic, it was shown that D-TCO judgements are not convergent with textbook based Judgements. With regard to

the *predictive validity* of the D-TCO measure, it can be concluded that student outcomes are to some extent related to D-TCO judgements of whether an item is taught. However since the size of the correlations varies between .37 and .46 at the test level (that is measured as the correlation between average student test score per teacher and percentage of ~taught~ items per teacher), and between .37 and .50 (in Vocational Education) and between .32 and .45 (in General Secondary Education) at the item level (that is, the correlation between average student score per item and percentage of teachers judging an item as 'taught'), it might be questioned whether the D-TCO measure is a good predictor for student achievement. A multiple regression analysis in which beside the TCO rating, the judgement of deviation of an item on certain characteristics was used as an extra independent variable did not improve the prediction of student achievement. Comparing the correlations we found with the correlation of the IEA-TCO measure with student scores collected in SIMS (test level: r=.22, (n=229); item level: Domestic Science Education: r=.32 (n=40), General Secondary Education: r= .00, n=40)), reported by Pelgrum (1990), showed that the correlations of the DTCO measure are higher at the test level and for General Secondary Education at the item level, but that these differences are not statistically significant. An analysis of the predictive validity of the H-TCO judgements showed that the correlations between these judgements and student achievement were higher than the correlations between student outcomes and D-TCO judgements. At a more specific level, it was found that, for those items that were judged as taught between the pretest and the posttest (according to the D-TCO measure), the increase of student scores was significantly higher than for items that were not taught between the pretest and the posttest, even when controlled for differences in student characteristics. An analysis of the influence of perceived difficulty of an item (measured by teacher estimations of percentage correct) showed that at the item level the D-TCO judgements as well as the H-TCO judgements are strongly related to the perceived difficulty. With regard to the *efficiency* of the D-TCO and the H-TCO instrument, it was shown that the time needed to judge one item with the D-TCO instrument is about 4 times longer than with the H-TCO instrument. The mean time needed to judge an item with the D-TCO instrument was 87 seconds compared to ~3 seconds for the H-TCO instrument.

References

Haan, D. M., de, (1992). Measuring test-curriculum overlap. Enschede: University of Twente.

Pelgrum, W. J. (1990). Educational Assessment: monitoring, evaluation and the curriculum. De Lier: Academisch Boeken Centrum.

INTEGRATING THE QUALITATIVE AND THE QUANTATATIVE:IN SEARCH OF A METHODOLOGY FOR COMPARING SCHOOLSYSTEM 'CASES'

S. Blom, Algemene Hogeschool Amsterdam, The Netherlands

By means of a comparison between a recent French and a recent Dutch national cohort, the 'Panel 1980' and the 'SLVO-cohort', some questions will be raised on how differences in the level of attainment in secondary education can be related to 'educational' characteristics of the system. As there are some clear and intriguing differences in level of educational attainment between the two mentioned countries (especially in the proportion of students that get into the academic stream and in the social composition of this stream) and as the system contexts seem, if anything, rather similar, one is tempted to look for explanations in different educational system characteristics.

However, questions arise, concerning firstly the comparability of the cohort data and secondly the supposed causal relationships between system characteristics and student outcomes. Thirdly, any general statements about this relationship on the mere basis of two 'cases', France and The Netherlands, seem hardly justifiable. on the other hand, it appears to be quite relevant nowadays to search for 'effective' system characteristics, just like have been found 'effective' school characteristics. Moreover, aren't national schoolsystems large enough units on their own, to allow some generalization? How can, in other words, a comparison between certain effects of two national schoolsystems lead to meaningful (i.e. more generalizable) and empirically justified explanations of the role of their 'educational' system characteristics?

In this contribution I will try to show how such a 'meaningful' comparison of two cases can be set up. This approach derives from a more extensive comparative study of French and Dutch secondary education, I am engaged in at this moment. The study will be published in 1993. The approach makes use of a theoretical (i.e. descriptive and analytical) framework that has been developed through former comparative studies in secondary education (Blom e.a. 1984; V08 e.a. 1985) and has been adapted, elaborated and legitimated through literaturestudy since the start of the comparative French-Dutch-schoolsystem study. The framework conceptualizes in a 'multi-level way' the relationship between educational characteristics and the intellectual (academic) level students reach in secondary education.

In order to clarify the ratio and background of the framework its 'history' will be followed as a chain of problemfinding and - solving, specifying other researchfindings and theoretical insights that have contributed to the most recent (but surely never final) version of the framework[1]. I will conclude by showing how the cohort comparison I started with, makes more sense while using the framework as a tool for assuring comparability and relating cohort data to educational characteristics.

1 In a French journal for educational research, "Perspectives Documentaries", this way of legitimating a approach is called an 'itinéraire de recherches'.

INTERNATIONALLY COMPARABLE INDICATORS OF EDUCATIONAL PROGRAMMES AND PROCESSES; IDENTIFICATION, MEASUREMENT AND INTERPRETATION

Jaap Scheerens, Department of Education, University of Twente, Enschede, The Netherlands

The conceptual framework of the OECD project on international educational indicators (INES project) distinguishes three domains: *demographic and economic background conditions, educational programmes and processes* and *educational outcomes*.

This paper focuses on a specific part of the second category: educational programmes and processes. Sub-categories of this domain that are being considered for precise definition and measurement within the INES project are:

- participation and flows within national educational systems;
- expectations and attitudes with respect to education of relevant audiences (general public, teachers, students, parents);
- resources applied (e.g. expenditures for elementary-secondary education per pupil);
- financial resources expended for key elements (e.g. per pupil expenditures for instructional materials);
- decision-making characteristics of educational systems;
- schooling processes;
- staff characteristics;
- instructional time;
- instructional context.

The part of the INES project that this paper deals with are the last four categories: schooling processes, staff characteristics, instructional time and instructional context.

The paper builds upon the set of indicators defined so far in the INES project within these four categories and an earlier publication of the author (Scheerens, 1990). In this latter publication the effective schools research was used as the major source for identifying "process indicators of school functioning". Another proposition was to distinguish between three levels of administrative use of process indicators: the national level, intermediate administrative levels (such as the municipality or the school district level) and the school level.

The present paper concentrates on the use of process indicators at the national level but uses an extended framework of organizational effectiveness to generate process indicators.

As encompassing framework of organizational effectiveness, proposed by Quinn and Rohrbaugh (1983) is employed as a conceptual structure from which relevant process indicators can be generated (see figure 1).

Human relations model

Quality of work life indicators
- respect
- participation in decision-making
- professional interaction
- performance feedback
- opportunity to use skills
- resources
- congruence personal/organizational goals

Internal process model

- planning documents
- disciplinary rules
- management information systems
- formalization of positions
- continuity in staffing & leadership
- integrated curricula
- attendance rates
- lessons "not given"

Open System model

- entrepreneurship
- collegiality
- capacity for self-evaluation and learning
- overt school marketing activities
- parental involvement
- boundary-spanning positions
- external change agents
- student enrollment figures
- resources (buildings, equipment)

Rational goal model

(school effectiveness research)
- educational leadership
- success-oriented ethos
- monitoring of students' progress
- time on task
- content-covered (opportunity to learn)

(broader set of educational goals)
- non-gradedness
- team teaching
- individualization, differentiation
- continuous learning route
- time spent on social, emotional, creative and moral development
- "learning to learn" activities
- diagnostic testing

Figure 1
Topics for process indicators generated from the Quinn and Rohrbaugh framework

Next, the question of measurement of these process indicators is taken up. Four practical measurement strategies are discerned: the use of available educational statistics, secondary analyses of existing data sets of internationally comparative surveys, "piggy-backing" ongoing or future international surveys and carrying out a specially designed "new" international survey as part of the INES project. Advantages and disadvantages of each of these approaches are discussed, the conclusion being that a periodic (every three to five years) parsimonious international survey included in the routine-work of national statistical agencies or inspectorates provides the most attractive measurement strategy, while interconnecting with ongoing or near-future international surveys is next best (while a combination of these, where the second approach could provide more in-depth information might be ideal).
A final problem that has to be faced is the evaluative interpretation of these process indicators. Particularly when causal links with educational outcomes have

not been firmly established in educational research (as is the case for many items from the more extended effectiveness framework in figure 1) the danger of "goal displacement" is very real. Some strategies to formulate operational standards to evaluate process indicators are discussed: estimates from empirical research about the increments in certain process variables that are needed in order to attain specific increments in outcomes, panels of educational experts that make judgements on values that should be reached on the process indicators and prespecified gains in "scores" on the process indicators over time. Particularly when the latter two approaches are chosen, panels should be composed of experts from all participating countries to arrive at internationally acceptable standards.

References

Quinn, R.E. & J. Rohrbaugh (1983). Spatial model of effectiveness criteria towards a competing values approach to organizational analysis. Management Science, 29, 363-377.

Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. School Effectiveness and School Improvement, 1(1), 61-80.

INTERNATIONAL SCHOOL EFFECTIVENESS IN RETROSPECT AND PROSPECT

Bert P.M. Creemers, Institute for Educational Research (RION), University of Groningen, Groningen, The Netherlands.

Introduction

In the United States there was a great deal of research in the seventies and eighties on the characteristics of effective schools in order to get indications how to improve schools. This resulted in the so-called five-factor-models, the most famous one being Edmonds five-factor model. Although in the last couple of years almost every presenter of research started with the confession that he does not believe in the five-factor-model anymore, this five-factor model has greatly influenced official research and practice in the United States.

Research in other countries could not provide evidence for this five and other factor models. Some effects were found in other studies as well, but others could not get empirical evidence.

In a junior high study by Mortimore, Sammons, Stoll, Lewis and Ecob (1988) twelve key-factors under the control of the head and the teachers were found which were crucial for effectiveness. Important factors in the United States, like high expectations, emphasis on basic skills, and a safe and orderly climate, do not appear in the same way in the studies by Mortimore et al. Furthermore, educational leadership gets a very specific content in the United Kingdom. The deputy head plays a major role in the effectiveness of schools. In the eighties also in the Netherlands studies were carried out with respect to school effectiveness. These studies came to the conclusion that between twelve and eighteen percent of the variance could be explained by school factors. The factors derived from

research in the United States did not explain much of the variance in the Dutch studies. Scheerens and Creemers provided an overview of Dutch studies so far (1989). In comparison with American and British research, it is striking that for educational leadership only in one empirical study evidence could be found. More empirical support was available for the factors orderly climate, frequent evaluation, and achievement orientation.
Critical analyses of studies in different countries have shown that we need improvements in theory and methodology of research on school effectiveness. Theoretically, we need conceptual models which should be multi-factor and multilevel. With respect to the methodology, the so-called outlier studies and survey research were heavily criticized and should be replaced by studies by means of observation and, ultimately, by experimentation, introducing effective school characteristics.

International studies
Because of the fact that the so-called school effectiveness factors/characteristics do not travel very well, international studies are required to obtain empirical evidence through countries with respect to school effectiveness in terms of learning outcomes and also with respect to the contextual differences between schools or school systems in different countries which could explain the differences in outcome. Using data sets from IEA studies, secondary analysis was carried out to look at common and uncommon effective schools factors (Lockheed and Komenan, 1989; Scheerens et al., 1989).
The principal conclusion of the Lockheed and Komenan study was that the achievement of students in both countries involved (Nigeria and Swaziland) was significantly affected by the school and classroom in which they were enrolled, once effects of family characteristics were controlled. However, the specific school and classroom level variables accounting for these differences were not the same in both countries. Effective teaching practices in one country setting could be entirely ineffective in another one. In the study, students in Nigerian mathematics classes who spent more time listening to the teacher introducing and reviewing mathematics outperformed those who were less exposed to direct instruction. The same result was not found in Swaziland. However, teaching time spent monitoring and evaluating student performance was positively associated with achievement in Swaziland, while it had no effect at all on achievement in Nigeria.
In the study by Scheerens et al. on the second maths study data seventeen countries were involved. The results of the secondary analysis suggest that a small number of school/class characteristics show a consistently positive association with mathematics achievement. These factors are positive expectations of pupil achievement and opportunity to learn. Other important school effectiveness factors, like teaching experience, time on task, frequent evaluation, were found to have weak and/or inconsistent effects. The analysis also supported the supposition that factors on the classroom level are somewhat stronger than factors on the school level.
Both studies point at the fact that international comparisons are important to develop a theory about educational effectiveness (as a combination of classroom effectiveness and factors on the school level creating conditions for effective classrooms). Available data sets of international studies are valuable sources for

secondary analyses in this field, however, to take the measures for factors on classroom and school level as they are or to create proxy variables. Most of the international studies are strong in output data but quite weak in process data, especially because it is impossible to collect data about the processes going on in schools and classrooms by means of observation. The theoretical framework of most of the studies is also not well-developed and especially in school effectiveness research a better conceptual framework is needed.

The International School Effectiveness Research Program (ISERP)

Based on the analysis of national and international studies the point of departure for the international research program could be formulated as follows:

- the conceptual framework;
- methodology of the study;
- the instruments used for the process and output measures;
- design of the studies (most of the studies are just a survey but not longitudinal and do not even have sound measures for output from the point of view of school effectiveness research);
- correction for intake measures used in the studies;
- the relationships between process data on the classroom, the school and the contextual level needed for interface between school, classroom and context.

The purpose of ISERP is to look at the communalities in school effectiveness factors in different countries and also in specific factors in countries that are connected with the societal educational system in these countries (for example, in the secondary analysis of Scheerens et al. some evidence could be found for the differences in school effectiveness in centralized and decentralized countries). So far ISERP has developed a conceptual model and has outlined a first study which will be presented.

References

Lockheed, M.E. & Komenan, A. (1989). Teaching Quality and Student Achievement in Africa: The Case of Nigeria and Swaziland. Great Britain: Pergamon Press plc.

Mortimore, P., Sammons, P., Stoll, L., Lewis, D. & Ecob, R. (1988). School Matters: The Junior Years. Somerset: Open Books.

Scheerens, J. et al. (1989). 'Generalizibility of Instructional and School Effectiveness Indicators across Nations'. In: B.P.M. Creemers & J. Scheerens (eds.), International Journal of Educational Research, Z (vol. 13), 789-801.

Scheerens, J. & Creemers, B.P.M. (1989). 'Conceptualizing School Effectiveness'. In: B.P.M. Creemers and J. Scheerens (eds.), International Journal of Educational Research, Z (vol.13), 691-707.

CONCEPTUAL AND FORMAL MODELS OF SCHOOL EFFECTIVENESS RESEARCH

Jaap Scheerens, Department of Education, University of Twente, Enschede, The Netherlands

School effectiveness as a field of educational inquiry has been characterized by a strongly empiricist orientation. Reviewers of the state of the art of school effectiveness research, for instance, Mortimore (1992), have pointed at the relative weakness of conceptualization and theory formation.

The models on school effectiveness that are presently available can be characterized as multilevel causal models of educational achievement in basic school subjects. There are three ways in which theoretical and conceptual work in school effectiveness can be advanced:

- connecting the empirical findings of school effectiveness research with more established theories in the social sciences, such as learning theory, cognitive psychology and public choice theory;
- carrying out empirical research that is aimed at hypotheses testing and model building rather than being a "fishing expedition" for positive correlates of achievement;
- bridging the gap between global and rough conceptual models on the one hand and precisely defined formal models (amenable to statistical analysis) on the other.

The symposium aims to contribute to this latter approach.

In the first paper, by Scheerens and Bosker, alternative causal specifications of the "nested layers" image of school functioning are examined. The statistical models that fit these alternative causal specifications are presented, together with a numerical example.

In the second paper, by Van Wieringen, the relationship between environmental administrative arrangements and school effectiveness is the central theme. The author presents a model in which "school autonomy" is a mediating factor between school-environmental arrangements and effectiveness enhancing school characteristics.

Giesbers, in the third paper, presents school organizational models, originating from E. Marx, in which the relationship between school organizational features and degrees of individualized instruction is the major issue. The relevance of these models for the conceptualization and research of school effectiveness is examined.

Obel will present the possibilities of Artificial Intelligence, in casu knowledge-based systems, to improve organizational theory. He formalizes organizational

contingency theory in order to build an expert system (the Organizational Consultant) for organizational diagnoses.
Bosker & Rijpkema present the design characteristics of a knowledge-based system that formalizes contingency theory for education. It is demonstrated how a formalization of this kind can help in theory formation.

CONCEPTUAL AND FORMAL MODELS OF SCHOOL EFFECTIVENESS

J. Scheerens, & R.J. Bosker, Department of Education, University of Twente, Enschede, The Netherlands

The following questions are addressed in this paper:

- is it useful to think in terms of a comprehensive multilevel school effectiveness model?
- what would be the key-variables of such a model?
- what are the basic explanatory mechanisms behind the interrelationship of factors within the model?
- which are competing causal specifications of the interrelationships?
- which statistical models are appropriate to compare competing causal specifications?

Most current reviews of the school effectiveness literature and proposals for school effectiveness models or even theories appear to accept Barr and Dreeben's basic notion of schools as "nested layers". According to this view the outcomes of one hierarchical level facilitate processes and outputs at the next.
So, for instance, managerial processes at school level are supposed to facilitate conditions of effective instruction at the classroom level. The point we which to make is that current conceptualizations of school effectiveness that roughly follow this framework are surrounded with considerable vagueness. In the process of attempts to formalize these models, some of this vagueness is likely to become evident and, in some cases, resolved by means of competing specifications.
First of all, the basic notion that conditions at high levels "somehow" facilitate conditions at lower levels can be interpreted in several ways; higher levels can be thought of as:

- providing contexts for lower level conditions made up of aggregates of lower level factors (contextual effects);
- minoring lower level conditions;
- serving as incentives of lower level conditions;
- material facilities for conditions at lower levels;
- overt measures to create effectiveness-enhancing conditions at lower levels (Scheerens, 1992).

Quite likely it will depend on the particular higher level variable which type of facilitating relationship is more plausible than another. The point is that conceptual models should be explicit on the particular interpretation of higher level facilitation.
Secondly, an even more basic issue is the question how comprehensive and complex our school effectiveness models should be. "Multilevel" is most likely to mean *four* levels: the level of the individual learner, the classroom level, the school

level and the immediate environment of the school (municipality, school district, local educational authority etc.).
At each level several relevant antecendent conditions of the ultimate criterion (student achievement) can be discerned. The set of relevant characteristics will become even larger when a broader conception of school organizational effectiveness is used. To avoid a prohibitive proliferation of variables, in our opinion, two mutually enforcing research strategies are to be chosen:

- to test global four-level models on large datasets, in which it is taken for granted that only summary indicators of key-factors are utilized;
- to test more specific two-level models in which more precise measurement of relevant variables is feasible.

In this paper the questions of key-variables and explanatory mechanisms will only be dealt with briefly, referring to other publications in which these issues are dealt with more extensively. We shall produce a list of key-variables as they have been identified in current reviews and mention some explanatory frameworks (didactic models, public choice theory, the cybernetic principle and models of coordination in educational organizations).
When turning to the question of alternative causal specifications within a global framework of schools as nested layers, we discern the following competing models:

additive vs interactive
According to additive models higher level conditions are seen as increments to variables operating at the lower level; e.g. achievement-oriented policy of an administrative level above the school "adds to" the effects of achievement-oriented policy at school level. In the interactive models higher level conditions impinge on the (causally interpreted) relationship between lower level antecendent conditions and the criterion variable; for instance as when instruction at classroom level is thought of as determining the impact of ability and effort of individual learners.
In terms of multilevel modelling the comparison of these two interpretations involves an interest in comparing intercepts (additive model) vs an interest in comparing slopes (interactive model).

recursive vs non-recursive
Negative correlations between variables that are taught to be effectiveness-enhancing and achievement are no exception in school effectiveness research. The inherent ambiguity in correlational research then allows for the interpretation that (e.g.) instructional processes are adapted to achievement levels. In fact, it is not at all implausible that several interrelationships among key-variables of school effectiveness models are in fact non-recursive. Questions about the recursiveness or non-recursiveness of certain interrelationships within school effectiveness models can be tackled in three ways:

- by means of experimental research;
- by means of alternative path-analytical models;
- by means of system-dynamic models.

System-dynamic approaches also bring the important question to the foreground as to which exogeneous factors can "break" repetitive cycles or feedback-loops. Conceptually this issue has to do with the primary "levers" of school effectiveness.

The idea of meta-feedback which originates from the image of the learning organization is also to be tackled by means of this methodology.

contextual vs "genuine" multilevel effects
A basic challenge of the nested-layers perspective on school functioning is the thesis that school effectiveness is largely determined by selection mechanisms (effective schools are schools that attract good students, good teachers and good administrators). "Higher level causation" according to this competing perspective would be largely determined by the contextual effects of aggregates (for instance: weak pupils do better in classes where average achievement is higher). Issues of contextual vs "genuine" multilevel effects can be settled by including both types of variables in multilevel models and by examining the relative magnitude of regression coefficients.

indirect vs direct causal effects
Conditions that are "more than one level up" with respect to educational achievement can either be seen as direct causes of achievement or as indirectly influencing achievement via intermediate levels. It should be noted that this sort of competing causal models cannot simply be settled by comparing different specifications of the usual LISREL-type or path-analytic models. Instead we would need multilevel path-analytic techniques, which are presently being developed.

additive vs synergetic interpretations
School effectiveness researchers, confronted with very low correlations between their antecedent conditions and achievement, have sometimes sought refrege in the thought that the joint effect of several variables, which individually appear to be of marginal influence, would "do the trick". The question is whether this magic of the whole being more than the seen of its parts is amenable to more precise and formal specification.
Mathematical formalization of each type of competing causal specification as well as a numerical example will be given in the paper.

References
Scheerens, J. (1991). School effectiveness: theory, research and practice. Cassell Publ. (in press).

INSTITUTIONAL CONTEXT, AUTONOMY AND QUALITY OF SCHOOLS

Fons van Wieringen, Department of Education, University of Amsterdam, Amsterdam, The Netherlands

Especially since the publication of Chubb & Moe's book on Politics, markets and America's schools in 1990 there is considerable debate whether the institutional context of schools does have a direct influence on the quality of schools. Chubb & Moe argue that an effective school organization cannot flourish without substantial school autonomy from direct external control. They make the point that the

bureaucratization of educational governance and administration has symply gone too far. Effective school organizations ask for a certain degree of school autonomy. This autonomy depends to a large degree on the institutional context of education - especially the privately v.s. publicly controlled aspects of the context of the school. "In the private sector, where schools are controlled by markets autonomy is generally high. In the public sector, where schools are controlled by politics autonomy is generally low." (p. 183). This paper discusses this relationship between institutional context (public/private sector), school autonomy and quality. The distribution of schools between the public and private sector shows a great deal of variance between different countries. To a large extent the market share of a private component in the educational sector is determined by national regulations and especially by national support for such a private sector. The first determines whether education is a state monopoly, the second determines the amount of public funds directed to the private sector. A comparative analyses shows a relationship between the volume of the private sector, the amount of public funds going into that sector and the density of state regulations. A complete private sector, as is discussed by Chubb & Moe, is almost completely free in setting its own regulations and in finding its own financial resources. A semi-private sector, as is the case in The Netherlands, derives its financial support almost completely from public funds and therefore have to accept a minimum amount of state regulations. The question is whether a private sector that is private in its main characteristics (private school-board, free selection of personnel and pupils) but performs public functions, attributes to its schools enough school autonomy? A school autonomy that is necessary in order to develop within the schools the characteristics of an effective school organization.

The research reported in this session.s consist of a survey research of a representative sample of public and private school-boards in The Netherlands. The main research questions concerns the relationship between the institutional (private/public) context and school autonomy. Do public and private school-board differ in their attributions of autonomy to schools? If there are such differences do they hold also when controlled for the scale of the school-boards itself? It might be that autonomy of schools is not so much a product of the private/public dimension of the institutional context but merely a consequence of the scale, the bureaucratization of the institutional context. This scale shows in The Netherlands enough differentiation within both private and public sectors.
The research project measured attribution of autonomy by school-boards to their school on five areas:

- selection, admission and dismission of pupils;
- selection of personnel;
- spending of educational current expenditures;
- spending of energy current expenditures;
- spending of material current expenditures.

As a general result it is found that the private sector grants more autonomy to their schools than the public sector. This seems to support the conclusions of Chubb & Moe.

We then analysed the differences controlled for the scale of the school-board. Especially in Dutch primary education there are considerable difference of scale: 70 % of the private school-boards have only one school to govern as compared to public education where this percentage is 30. In the public sector the scale of operations is widened, this might cause a more bureaucratic treatment of the public schools.
Analyses shows that for the group of school-boards (public and private) with one school there are only small differences between private and public school-boards. These differences are much larger within the group of school-boards with more than one school to govern. Private boards in this group do attach much more autonomy to their school, than public boards. Scale of the board in terms of numbers of governed schools does show a different relationship with attribution of autonomy in the private as compared to the public sector.
The paper finally discusses some of the possible explanations for the effect of scale in both sectors.
Finally the paper discusses the research model used, by comparing the model of Chubb & Moe with the model used here and with other possible models. These research models can be seen in context of underlying causal models of effective schools. Two of these models are discussed as far as their conceptual clearity concerns and as far as their technical-instrumental problems in measuring the stated concepts concerns.

References

Chubb, J.E. & Moe, T.M. (1990). Politics, Markets, and America's Schools. Washington, D.C.: The Brookings Institution.

THE MARX-MODELS AS CONCEPTUAL MODELS IN SCHOOL EFFECTIVENESS RESEARCH

J. Giesbers, Catholic University Nijmegen, Nijmegen, The Netherlands

Introduction

School effectiveness research is research into the actual or supposed relationships between independent variables in terms of school characteristics, however defined, and dependent variables in terms of effects on the individual-, class- or schoollevel.
Many conceptual models can be used as searchlights in this research (see Scheerens, 1990). This paper focuses on one set, viz. the Marx-models. Their first conceptualization and formulation date back to 1975, well before the effectiveness-movement got into full swing. Over the years they have been developed and refined, both theoretically and empirically, though the basic tenets and fundamental ideas have remained the same. The thesis of this paper is that the Marx-theory can provide a theoretical basis for the research into questions pertaining to what makes school effective, how and why. Putting if differently: this

paper considers the Marx-models as school effectiveness models which focus on the contextually determined effectiveness on the meso level of the school.

School effectiveness research

This research has been and is being criticized on methodological and theoretical or conceptual grounds. According to Rosenholtz: "Careful analysts of the effective schools research find it not without methodological problems" (1985, p. 353): the comparison of extreme outliers while neglecting average schools and the measurement of random error - the inaccurate use of the case-studymethod - the use of cross-sectional instead of longitudinal designs - the insufficient control for confounding variables e.g. pupil SES - the generalizability to primary schools only. For a useful overview, see Scheerens 1990.

This research, however, also came under fire because of theoretical or conceptual problems. There is no simple, unidimensional measure of school effectiveness (Mortimore, 1991) and even the definitions of this key concept vary. Some researchers consider schools effective when their students have sufficient scores on tests of basic skills. For other researchers, this cognitive and instrumental conceptualization and measure is far too narrow and they prefer, beside cognitive, also affective and behavioural interpretations and measures. This paper does not focus on micro effectiveness on the pupil-level but concentrates on schoolorganizational effectiveness on the meso level.

In view of the differing interpretations and measures of schooleffectiveness, it is not to be wondered at that the research resembles a patchwork quilt and that the theoretical status, in spite of a growing number of promising attempts to raise it, is not very high. We know a good deal of what works in education; we have as yet insufficient insight into the why and how. Because of this theoretical poverty of school effectiveness research (Scheerens and Creemers, 1989), those concerned should focus on the generation of theory (Mortimore, 1991) and should use any theoretical instrument available to interpret the results of their empirical work (Scheerens and Stoel, 1988).

This paper presents the Marx-theory as a theoretical underpinning of school effectiveness research. His theoretical position concerns the relationship between the goals of schools, the structure of schoolorganizations, and the context which influences goals and structures. Putting it differently: dependent on the context, some organizational structures are better suited in terms of effectiveness for the realization of educational and instructional goals than others. Because of this central thesis of his work and because of another criterion of effectiveness, viz. the consistency between the educational and organizational models, the Marx-models can be seen as school effectiveness models on the meso level and as instruments in school effectiveness research.

The Marx-models

In 1975 a doctoral dissertation appeared, entitled The Organization of Comprehensive Schools from an Educational Point of View. The study presented five educational-organizational models of school.

Each of Marx's models consists of an educational and an organizational system. These systems are mutually consistent "in the sense that the organization is a suitable means for the effective functioning of the educational model" (Marx, 1975,

p. 33). The educational and organizational systems of a model are therefore interrelated on an end-means basis, in which the organization of the comprehensive school is the means to make a desirable educational system function effectively. In all, Marx distinguishes five mutually consistent educational and organizational models (1. The vertical-segmental; 2. the restricted vertical line and staff; 3. the horizontal-collegial; 4. the integrated-differentiated; 5. the plural-process model).

The educational models, each consisting of seven components, describe the order of teaching-learning situations. The distinction between the educational models is founded on the individualizing capacity of schools, that is to say the extent to which the school allows students to partake of education geared towards making maximum allowance for a student's individual capacities.

In the organizational models a number of persons and sections are distinguished (management, departments, educational counsellors, consultative bodies). The functions of these persons and sections relevant to the realization of the educational model are described, as well as their mutual interrelationships with respect to processes of policy-making and policy implementation.

The various consistent educational and organizational models are arranged by Marx in a sequential series, which he terms a "structural development model". A structural development model can be said to exist if the consistent educational and organizational models can be fitted into a series in such a way that their positions are indications of the positions schools must pass through in a process of systematic change. The constructed educational-organizational models constitute an ascending series within this structural development model, in which each subsequent model is more complex than its predecessor in educational-organizational respect. The most complex educational-organizational model is better suited to realizing more complex educational goals than any other preceding model in the series.

It is characteristic of the series of models and the structural development model, that the administrative and instructional domains are increasingly interwoven as the models acquire greater educational and organizational complexity. In the more complex models teachers are to a relatively large extent involved in policy decisions, the management gains importance as a (co-)initiator and co-ordinator of educational policy.

The progressive interrelation of the administrative and instructional domains in Marx's series of models does not exclusively matter to the realization of educational goals of ever-increasing complexity. Along with the progressive interrrelation of the administrative and instructional domains, the school's capacity to respond to changing circumstances increases as well.

This latter evolution is highly important. Policy formation particularly refers to decision-making with respect to desirable educational changes and innovations, and the provisions which are needed. The respective models organize these processes of policy formation differently. The more complex models allow for more complex policy formation. Marx's series of models in general and his notion of the interrelation of the administrative and instructional domains in particular, are therefore important to processes of policy-making and organization development in schools and to raising their effectiveness.

The Marx-models as school effectiveness models

In his contribution to the 1989 Dutch Educational Research Conference, Scheerens compared two substreams in the field of research into educational organizations, viz. schoolorganizational effectiveness and the Marx-models. He described the differences and pleads for a fruitful interplay between these two substreams. In this section of my paper I intend to further elaborate this plea. Though it is certainly true that there are differences between the Marx-models and other school effectiveness models, it is also true that the differences do not prevent the Marx-theory and models from contributing fruitfully to school effectiveness research. To illustrate and defend my thesis that the Marx-models may be conceptualized as school effectiveness models on the meso level, I shall use his valedictory address as professor of organizational science of Leyden University. In this address he reformulated and condensed the original educational-organizational models presented in his doctoral dissertation of 1975. In 1990 he distinguishes the following organizational models:

- The segmental organisation, (in 1975: model 1) effective (and efficient) in a subject oriented and restricted professional (Hoyle 1975) sense;
- The cooperative regulated organisation, (in 1975: models 2 and 3), effective (and responsive, Van Wieringen, 1989) not only in a subject-oriented, restricted professional but additionally in a wider educational and extended professional (Hoyle 1975) sense;
- The learning organisation (in 1975: models 4 and 5), which is perhaps less effective than the first two models in the sense of realization of given goals but which achieves effectiveness because it develops and innovates the given goals into new goals which better answer contextual demands. In this model effectiveness is synonymous with flexible innovativeness and creativity.

In my paper I shall further describe and develop these three models and their use as theoretical searchlights in school effectiveness research. As they have not been tested empirically since their first formulation in 1990, case studies followed by larger scale surveys are called for. Special attention will be paid to Marx's changing definitions and conceptualizations of effectiveness in which contextual conditions play on important part. An attempt will be made to link his views with Quinn and Rohrbaugh's fourfold conceptualization of organizational effectiveness (Schaveling, 1989).

References

Hoyle, E. (1975). Leadership and decision-making in education. In M. Hughes, (Ed.), Administering education: international challenge. London.

Marx, E., (1975). De organisatie van scholengemeenschappen in onderwijskundige optiek. Groningen: Wolters-Noordhoff.

Marx, E., (1990). Schoolontwikkeling en milieugericht onderwijs. Alphen a/d Rijn: Samsom.

Mortimore, P., (1991). School Effectiveness Research: Which Way at the Crossroads? In: School Effectiveness and School Improvement, 2, nr. 3, 213-229.

Rosenholtz, S., (1985). Effective Schools: Interpreting the Evidence. American Journal of Education, 93, 352-388.

Schaveling, J., (1989). Effectiviteit: ook een trend in het onderwijs, Meso, 44, 2-7.

Scheerens, J., (1989). Onderzoek naar schoolorganisatorische effectiviteit en de contingentie benadering. In J. Scheerens, & J. Verhoeven (red.), Schoolorganisatie, beleid en onderwijskwaliteit. Amsterdam/Lisse: Swets & Zeitlinger.
Scheerens, J., (1990). Wat maakt scholen effectief? 's-Gravenhage: SVO.
Scheerens, J., & Creemers, B., (1989). Conceptualizing School Effectiveness. International Journal of Educational Research, 691-705.
Scheerens, J., & Stoel, W. (red.) (1988). Effectiviteit van onderwijsorganisaties. Lisse: Swets & Zeitlinger.
Wieringen, A. van, (1989). Bestuur en management van onderwijsinstellingen. Groningen: Wolters-Noordhoff.

TEACHING ORGANIZATION THEORY AND DESIGN USING AN EXPERT SYSTEM

R.M. Burton, The Fuqua School of Business, Duke University, Durham, N.C.; B. Obel, Department of Management, Odense University, Denmark

Organization theory is a large body of knowledge of loosely related language, definitions, hypotheses, analyses and conclusions. Organization theory is multi-disciplinary where the separate disciplines have created the questions, hypotheses, methodologies and conclusions - each built on related yet distinct views of the world. Organization theory is a positive science which focuses on an understanding of organizations.

Organizational design is a normative science which focuses on the creation of organization to obtain given goals. Of course, organizational design must incorporate the knowledge from organization theory; otherwise, it is an exercise in abstraction without empirical content.

There is a very large knowledge base in the literature and common knowledge where both academics and practitioners are experts. However, this vast knowledge is fragmented in many different theoretical constructs and practical knowledge. There is not a single coherent structure for our knowledge on organizational design. The beginnings of such an effort is the consideration of normative recommendations as well as efficiency and effectiveness hypotheses from the literature.

The contingencies for organizational design potential include strategy, size and power control as well as environment and technology (Robbins 1990, p. 22). Textbooks on organizational theory and design usually present each contingency and its effects on the organizational design separately. The student is then left with the complicated task of combining the different effects.

Organizational Consultant is an expert system that can be used in teaching organization theory and design. It is designed to teach the student the relationships between each of the contingencies and the organization design as well as balancing the various factors.

Despite the diversity and vastness of the literature, contingency theory is the dominant theme in organization theory (Robbins, 1990; Daft 1986). Size, strategy, uncertainty, environment, power have all been offered and argued as imperatives to determine the organization's structure. The controversy over which organizational imperative is the correct one has largely disappeared. The more general contingency concept would suggest that each factor is important in structuring the organization. The organization must fit with one or several of these determining factors in order to be efficient, effective, or at least viable. In Figure 1, these contingency concepts are summarized. For the fit criteria, the contingency factors lead to an appropriate organization structure and properties.

The *Organizational Consultant* expert system analyzes the current organizational structure by queering the user about facts related to the functioning of the organization. The structure is then described in the terms that are later used to describe the recommended structure.

Next the expert system asks for input related to the major situational variables. Based on this input the system recommends the most likely structure that gives the best fit with the specified situation. The situation itself is analyzed and possible situational misfits are given. Finally, the current and prescribed organizational structure is compared and possible changes recommended. It also provides a way to perform sensitivity analysis. The analysis consists of an analysis of size, management style, strategy and organization structure. Each can be run separably. The results from the subanalysis of size, strategy and management style are input to the organizational analysis.

The contingency model of organizational theory

The Contingency Factors for Organizational Structure:

Strategy
Size
Technology
Environment
Power Desires

Fit Criteria:

Effectiveness
Efficiency
Viability

Properties and Structure of the Organization:

Structure: Simple, functional, divisional, machine bureaucracy, matrix, etc.

Properties:

Complexity and differentiation
Centralization
Span of Control
Rules
Procedures
Professionalization
Activities

Meetings
Reports
Communications

Organizational Consultant provides context sensitive help. For each of the concepts used in organization theory and design, see e.g. Fig. 1, clear and precise definitions are provided with some explanation why the factor is important for organizational design.

The results from *Organizational Consultant* include an explanation written in plain English for each conclusion that is given. These explanations present the piece of the organization theory that was used to obtain the conclusion for the particular case. Thus the organization theory can be shown and explained to the students by letting them run a number of well selected case studies. It presents an interactive way to teach organization theory. This is further enhanced by easy features to run sensitivity analysis which shows the students the effect of marginal changes in the contingencies. The *Organizational Consultant* has been used as a teaching support tool for teaching MBA students, executive MBA students, and executives. It is our experience that such a tool can be an effective support that provides the students with greater insights in a shorter period of time than can be obtained by more traditional teaching methods.

It has been proved to be particularly effective in teaching classes with a large number of students where it provides a feed-back to the students in a way that is normally not possible. Additionally, it teaches the students to be precise and systematical in analyzing organizational cases.

For each case the following process is recommended:

- *Enter input;*.Interpret the information given in the case and note issues on which answers are difficult - possible candidate for a sensitivity analysis.
- *Run sub-analyses on management style, size and strategy;*.Carefully review. Change results if necessarily. Note the change as a possible candidate for a sensitivity analysis.
- *Run organizational analysis:* Carefully review recommendations *and* explanations. Note the certainty factor to see how confident the Organizational Consultant is. Carefully review the situational misfits. If there are misfits, can or should they be removed? If so, how? Note as a possible candidate for a sensitivity analysis.
- *Run sensitivity analyses;* Go through the previous steps for an appropriate number of changes. Stop when satisfied with the results.
- *Review the detailed recommendations;* Evaluate if they are appropriate for the case.
- *Interpret what the recommendations mean for the particular case;* Make the recommendation specific to the case. Use the vocabulary of the case.
- *Review the change process;* Create a plan for implementing the changes.

References

Baligh, H.H., R.M. Burton, & B. Obel. (1990). Devising Expert Systems in Organization Theory: The Organizational Consultant In M. Masuch (ed.) Organization, Management and Expert Systems. de Gruyter.

Robbins, S. A. (1990). Organizational Design: Structure, Design and Applications. Prentice Hall.

CHAPTER 2

CURRICULUM

INTRODUCTION TO THEME 2: CURRICULUM

Jan van den Akker, Uwe Hameyer, & Jan Terwel

As coordinators of the Curriculum theme (Jan van den Akker, Uwe Hameyer, and Jan Terwel) we have been pleased with the wide response on our invitation for proposals. After having reviewed the proposals we can announce an interesting program that includes 6 symposia (with 24 presentations), and 30 individual paper presentations. This number, and the fact that the presenters come from fourteen different countries, reflects the broad interest among curriculum researchers in Europe to engage in an scholarly exchange of ideas and findings. It is encouraging that so many researchers with various backgrounds and specializations in the curriculum field are willing to participate in this first opportunity to meet on a European scale. A special session will be held at the conference to explore the possibility of creating an European Network for Curriculum Researchers that may facilitate the organization of future conferences in the curriculum domain.

The scope of the presentations is very broad, as one might expect in the curriculum field, covering may educational levels, settings, subjects and products. This variety is well illustrated by the titles of the different symposia:

- "Searching for more effective curriculum development strategies".
- "Values in education in Europe: A CIDREE collaborative programme".
- "Curriculum reseach and curriculum productivity".
- "Some results from IEA's study 'Computers in Education'".
- "Evaluating the implementation of English, mathematics and science in the English national curiculum: Procedures and early findings".
- "Courseware in the curriculum: Demands on teachers".

A program highlight will be the keynote address by Decker Walker about 'formative research', a kind of research intended to inform the development of curriculum materials.

It promises to be an address that will raise interesting discussions on both theoretical, methodological, as practical issues.

We also expect a somewhat provocative presentation by Ian Westbury, who will discuss European curriculum research from his American perspective.

But of course many other interesting presentations will be delivered, e.g. in the paper sessions, your choice depending on your personal preferences.

Anyway we see this conference offer as an encouraging start for improving the communication between curriculum researchers in Europe (as a matter of course without excluding participation of non-European colleagues). Let us hope it will evolve as a new and long tradition!

EFFECTS OF COOPERATIVE LEARNING AND GROUP COMPOSITION IN A SECONDARY MATHEMATICS CURRICULUM

P. van den Eeden, Department of Social Science Methodology, Free University Amsterdam, Amsterdam; T. Mooij, Institute for Applied Social Sciences, University of Nijmegen, Nijmegen, & J. Terwel, Graduate School of Teaching and Learning, University of Amsterdam, Amsterdam, The Netherlands

Research problem

A great deal of research has been carried out into the effects of learning in small groups. As a result, some consensus has developed in education circles about the relevance of learning in small classroom groups. In addition, investigations have been conducted into the effects of class characteristics on the learning processes and outcomes. It therefore seems worthwhile to combine both types of analyses. The research problem addressed in this paper is in line with both developments. The question is whether or not (a) there are differences between subgroups of students within classes in the effects of learning mathematics and (b) if so, whether those differences exist across classes. To put it differently, which level is more important, the subgroup level or the class level in explaining the effects of learning?

Curriculum characteristics

The curriculum for cooperative learning was implemented in four schools, including 'integrated schools' and middle schools. The curriculum was confined to second-grade secondary mathematics education. All schools used the same textbook, entitled 'Wiskunde Lijn' (published by Jacob Dijkstra in Groningen). This textbook was adapted to a curriculum for cooperative learning according to the AGO model (AGO is a Dutch acronym for 'adaptive group teaching'). This model combines 'adaptive teaching' and 'cooperative learning'. (Herfs, Mertens, Perrenet & Terwel 1991).

The various instructional strategies in the AGO model are designed to develop higher order thinking. The AGO model tries to create a balance between (i) basic skills and concepts developed through whole-class instruction and (ii) problem solving developed by guided-discovery learning in cooperative groups. In the

model arrangements are made to adapt teaching to differences in aptitude between students.

Theory, research problem and variables

Learning in small groups has received much attention from practitioners and researchers. The last decade there has been a growing movement in cooperative learning all over the world. Many researchers, curriculum specialists and teacher educators are involved in cooperative learning. Research into cooperative learning can seen in the light of two theoretical and empirical traditions.

First there are studies at the classroom level. These studies are directed to the implementation and effects of certain class characteristics. Among these characteristics are composition of classes and instructional strategies (Dar & Resh 1986, Slavin 1989, Terwel & Van den Eeden 1990, Mooij 1987). The concept of the class socio-learning environment (SLE) (see Dar & Resh, 1986) is relevant to our investigation. Effects of class-composition can be explained by differences in the quality of the SLE (rich versus poor SLEs). Class composition can be described in terms of the ratio of students with rich personal resources versus poorly resourced students.

A second research tradition is directed at the small group level e.g. effects of instructional characteristics and small group composition. This research was carried out by Webb (1982), Webb & Kenderski (1984), Hallinan (1987) and Leechor (1988).

In explaining the findings of these small-group studies various factors emerge. Pupils in small groups are confronted by their fellow-pupils in the group with different solutions and points of view. This may lead to socio-cognitive conflicts, which are accompanied by feelings of uncertainty. This may cause a willingness in pupils to reconsider their own solutions from a different perspective. The resulting processes may stimulate higher cognitive skills.

Small groups offer group members the opportunity to profit from the knowledge available in the group as a collective. This may take the form of knowledge, skills or experiences which are not possessed by each member of three group. Pupils use each other as 'resources' under these circumstances. Leechor (1988) calls this 'resource sharing'.

Collaboration in small groups means that pupils are stimulated to verbalize their thoughts. Such verbalizations facilitate understanding through cognitive reorganization on the principle that 'He who teaches learns the most'. Offering and receiving explanations enhances the learning process. Group members do not only profit from the knowledge and insights transmitted through 'peer tutoring', but they can also internalize effective problem solving strategies by participating in the collective solution procedures.

Research in which both traditions (classroom research and small- group research) are combined does almost not exist. Until now effects of learning have been attributed either to variables at the class level or to small-group variables. In both traditions effects have been found, but no indications are available about the relative impact of class variables and small-group variables on achievement. From a multilevel perspective analyses can be carried out in which both types of effect (class and small group) can be studied simultaneously. The research problem can be specified as follows:

What are the effects of instructional and composition variables at the class and subgroup levels on the learning of mathematics?.

Data, method and analysis

The analysis was carried out by means of a data set taken from the AGO project for 12 to 16 year-olds. In total 381 students, 100 subgroups, and 14 classes were used in the analysis. The study contains, among other variables, a pretest, observations of the cooperative learning process, and a posttest. The student pretest is a test in mathematical reasoning. The posttest is a curriculum-specific test in mathematics. In addition, the attitudes of the students were measured by means of two questionnaires. Each of these was administered before and after implementation of the new curriculum. For the class and subgroup some compositional characteristics were computed. The observations were conducted by means of a recently developed instrument called AIOS (Adaptive Instruction Observation Scale), a systematic observation system with time sampling. The observations were aimed at instructional characteristics. In the analysis the random coefficient model of multilevel analysis was used. Three levels are involved: the student, the subgroup and the class level.

Outcome and conclusion

Three separate analyses were carried out with 'achievement in mathematics', 'attitude toward mathematics' and 'attitude toward the teacher' as dependent variables. From an initial analysis it appeared as if there were significant differences in regressions (both in intercepts and in slopes) between the subgroups, but after introduction of the class level these regressions disappeared. This outcome means that there is some within-class homogeneity in the subgroup effects. It also shows that in this investigation no autonomous role is played by the subgroup, due to the crucial role of the class. Consequential the general conclusion is that the class level is more important than the small-group level in explaining the effects of learning. The specific conclusions are:

- achievement in mathematics is positively related to Aptitude in mathematics and Attitude towards mathematics at the individual level;
- achievement in mathematics is positively related to instructional (groupwork) and compositional (aptitude) variables at the class level.

Discussion

This conclusion sheds an unexpected light on the theoretical and empirical relevance of instructional subgroups (the level of small groups within classes). At least it suggests the need to incorporate the class in a theory of subgroup effects in education. There are two possible explanations for the absence of differences in effects of learning related to the subgroup level:

- subgroups within a class do not differ in composition from each other, each subgroup in a class being more or less a reflection of the class as a whole;
- the absence of differences in effects between subgroups may be caused by the specific model of instruction. In the AGO model groupwork takes less than one third of the time spent in class. All group activities are embedded in procedures in the class as a whole. Resource sharing between subgroups in a class is a

characteristic of the model. Probably this exchange of information between subgroups diminishes the differences between subgroupes.
The conclusions regarding the specific variables at the individual and class level are in line with theoretical and other research. The positive effects of groupwork and class composition are important for the theory and practice of cooperative learning strategies.

References

Dar, Y. & Resh, N. (1986). Classroom Composition and Pupil Achievement: A study of the Effects of Ability-Based Classes. New York: Gordon & Breach.

Hallinan, M.T. (1987). Ability Grouping and Student Learning. In: M.T. Hallinan (Ed.), The Social Organization of Schools, pp. 41-69. New York: Plenum.

Herfs, P.G.P., Mertens, E.H.M., J. Chr.Perrenet & J. Terwel (1991). Leren door samenwerken. Amsterdam/Lisse: Swets & Zeitlinger (SVO Forumreeks).

Leechor, C. (1988). How High and Low Achieving Students Differentially Benefit From Working Together in Cooperative Small Groups. Stanford: Stanford University, School of Education (Dissertatie).

Mooij, T. (1987). Interactional Multi-level Investigation into Pupil Behaviour, Achievement, Competence, and Orientation in Educational situations. 's Gravenhage/Nijmegen: SVO/ITS.

Slavin, R.E. (1989). Cooperative Learning and Student Achievement. In: R.E. Slavin, School and Classroom Organisation, pp. 129-156. Hillsdale, N.J.: Erlbaum.

Terwel, J. & Van den Eeden, P. (1990). Effecten van gedifferentieerd wiskunde onderwijs: de toepassing van een model voor multilevelanalyse bij curriculumevaluatie. Tijdschrift voor Onderwijsresearch, 15, 5, 273-284.

Webb, N.M. (1982) Group Composition, Group Interaction, and Achievement in Cooperative Small Groups. Journal of Educational Psychology,74. 475-484.

Webb, N.M. & Kenderski (1984). Student interaction and learning in small group and whole class settings. In: P.L. Peterson, L.C. Wilkinson & M. Hallinan (eds): The Social Context of Instruction. pp. 153-170. Orlando: Academic Press.

Webb, N.M. (1988). Peer Interaction and Learning in Small Groups. Paper presented at the Annual Meeting of the American Educational Research Association. New Orleans.

MATHEMATICS TEXTBOOKS AND LEARNING RESULTS IN PRIMARY SCHOOL GRADES 1 THROUGH 3

Marja van den Heuvel-Panhuizen, & Koeno Gravemeijer, Freudenthal Institute, University of Utrecht, Utrecht, The Netherlands

Introduction to the problem

The aim of the 1987 started MORE project 1 was to find out to what degree textbooks and certain teacher characteristics influence education and how this influence carries over into learning results. Two different domain specific instruction theories, the mechanistic theory and the realistic theory, were central to this research (Gravemeijer et al., 1991). The realistic approach can be described by means of the level theory of Van Hiele (1973), the didactical phenomenology of Freudenthal (1983), and the principle of progressive schematization developed by the Wiskobas group (Treffers, 1987). The characteristics of the mechanistic approach resembles the task analytic instruction theory (cf. a.o. Gagné, 1977, and Resnick & Ford, 1981). Apart from these domain specific instruction theories, also theories of curriculum implementation are taken into account. Especially the concept of idea consistent adaptation (Gravemeijer, 1987) - which is an elaboration of the ideas of Fullan & Pomfret (1977) - was paramount. As a consequence the beliefs of the teachers (Thompson, 1984; Fullan 1983) were also important.

The research was divided into a number of components. This paper only deals with one of them, i.e. the influence of textbooks on performance in mathematics. In particular, it addresses the question of whether it matters if pupils are taught on the basis of mechanistic or realistic textbooks. Former research did not give an unambiguous answer to this question. The PPON-research (Wijnstra, 1988), for instance, revealed a positive effect of the realistic textbooks, whereas Harskamp & Suhre (1986) came to the conclusion that textbooks do not matter.

Research method

Eight schools (8-9 classes) which work with the mechanistic textbook series 'Towards independently working in Arithmetic' (NZR) and ten schools (13-15 classes) which work with the realistic textbook series 'The world in numbers' (WIG) participated in the research. Both groups of schools had a heterogeneous composition and involved, for instance, both innercity schools and rural schools. In total about 430 pupils were examined from grades 1 through 3. Classroom tests were developed to measure performance in mathematics. Each year group was

measured four times. The test battery consisted of ten general mathematics tests (including two repeat tests) and two number facts tests. The general test comprises a broad range of problems including context problems as well as bare sums, whereas the number facts tests only contains bare sums, which have to be solved within a restricted period of time. In order to be able to provide as reliable an answer as possible to the question of effect, also some background data have been collected about the pupils and the general teaching conditions. As regards the pupils, data have been collected concerning their intelligence, social class level and beginning math level. Data about the general teaching conditions concerned the time spent on teaching mathematics, the task-orientation, the form of classroom organization, and the degree to which the textbook was used in the lessons. The analysis was carried out at pupil level and including total scores as well as changes in scores for five subsidiary skills. In addition, analyses of specific numerical operations were carried out for each year group. All these analyses involved the use of covariance analyses, with the pupil characteristics intelligence, beginning math level and social class level as covariables. The two groups of pupils showed no difference as regards the general teaching conditions.

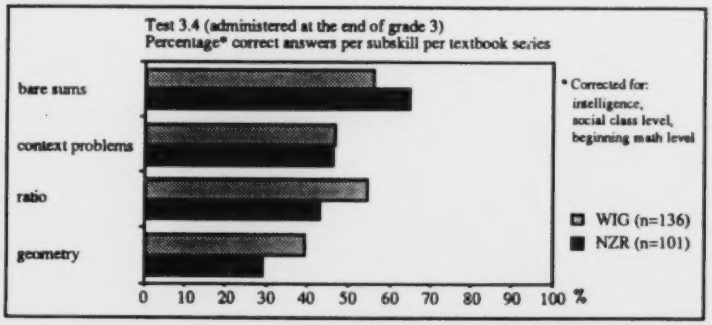
Results

The textbook analysis showed that there are great differences between the mechanistic NZR and the realistic WiG. The differences not only concern the expected differences of the didactical approach, but also the learning content. NZR almost exclusively concentrates on arithmetic, whereas WIG contains other branches of mathematics such as geometry and ratio. There are also differences between the two textbook series as regards the shared component of arithmetic. An additional difference lies in the field of applications. In NZR these are all but non-existent. Another, rather surprising fact is that the two textbook series are not particularly homogeneous over the three grades.

The analyses of test results showed that, seen over the three years, the NZR pupils are better in bare sums, whereas the WIG pupils are better in the subsidiary skills of geometry and ratio, as well as are more skilled with numbers at the beginning.2 For the context problems (which included the ratio problems) the scores vary. Presumably basic skills in arithmetic tip the balance at one time and familiarity with applications at another time.

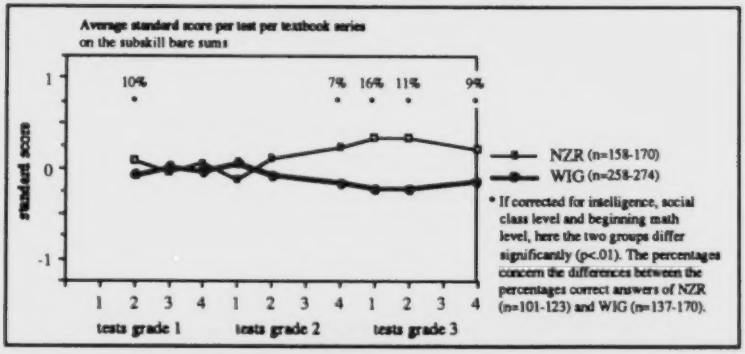
This overall outcome also emerged in the skills profile at the end of grade 3 (figure 1). The scores of the NZR pupils (corrected for intelligence, beginning math level and social class level) are higher for the bare sums, whereas those of the WIG pupils are higher in the subsidiary skills of geometry and ratio. The size of the differences is around ten percent. With respect to the context problems the scores do not differ.

Fig. 1


Test 3.4 (administered at the end of grade 3)
Percentage* correct answers per subskill per textbook series
bare sums
context problems
ratio
geometry
0 10 20 30 40 50 60 70 80 90 100 %
* Corrected for: intelligence, social class level, beginning math level
WIG (n=136)
NZR (n=101)

The presented learning content is also recognizable in the scores during the three years for each subskillary skill. Figure 2 shows the successive scores on the subskill bare sums. According to the earlier start in NZR with bare sums the NZR pupils score better at the first test which includes this subskill. The WIG pupils do not learn these sums until the end of grade 1. At that moment the difference decreases, but only for a while. At the end of grade 2 the NZR pupils have again better scores and this is continued in grade 3. Only at the end of it there seems to be a turn of the tide.

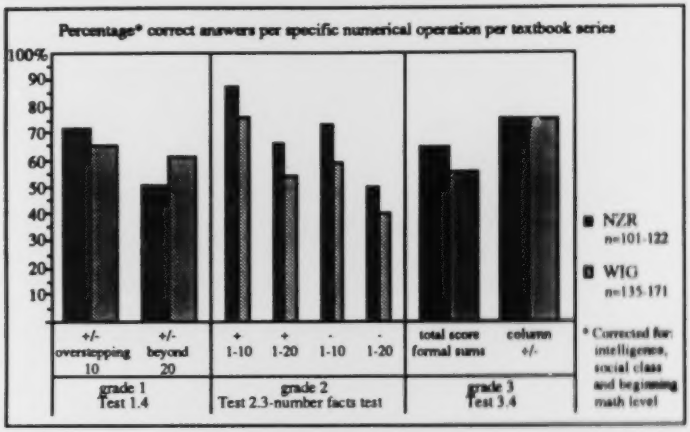
Fig. 2


Average standard score per test per textbook series
on the subskill bare sums
standard score
1
0
-1
10%
7% 16% 11%
9%
1 2 3 4 1 2 3 4 1 2 3 4
tests grade 1
tests grade 2
tests grade 3
NZR (n=158-170)
WIG (n=258-274)
* If corrected for intelligence, social class level and beginning math level, here the two groups differ significantly (p<.01). The percentages concern the differences between the percentages correct answers of NZR (n=101-123) and WIG (n=137-170).

The analyses of the specific numerical operations revealed that even from a closer point of view the scores are in accordance with the presented learning content. The NZR pupils are better in sums overstepping 10 and the WIG pupils in sums beyond 20. The NZR pupils obtain higher scores on the number facts tests in grade 2 (and also in grade 3). The NZR and the WIG pupils perform equally well in column addition and subtraction in grade 3 (figure 3). The latter could be an

indication that not only the quantity of presented learning content matters but also the chosen didactical approach.

Fig.3



Conclusions

The first conclusion to be drawn is that the two textbook series under investigation are non-homogeneous and different from each other. The second conclusion is that the test results clearly indicate that textbooks matter.

Discussion

The results in this research indicate that textbooks can be a powerful instrument in shaping learning results. The analysis of the teaching protocols (leave out of consideration in this paper), however, shows that the implementation of realistic mathematics instruction is still far from ideal. To put it differently, the limits of what is attainable have certainly not been reached yet.

References

Freudenthal, H. (1983). Didactical Phenomenology of Mathematical Structures. Dordrecht: Reidel.

Fullan M. & A. Pomfret (1977). Research on Curriculum and Instruction Implementation. Review of Educational Research, 47 (2), 335-397.

Fullan, M. (1983). The Meaning of Educational Change: a Synopsis. Pedagogisch Tijdschrift/Forum voor Opvoedkunde, 9, 454-464.

Gagné, R. M. (1977). The Conditions of Learning (3rd ed.). New York.

Gravemeijer, K. (1987). Innoveren en experimenteren. In E. Feijs & E. de Moor (eds.). Panamacursusboek 5. Innovatie realistisch reken-wiskundeonderwijs. Utrecht: SOL/OW&OC.

Gravemeijer, K., M. v.d. Heuvel-Panhuizen, G. v. Donselaar, N. Ruesink, W. Vermeulen, L. Streefland, E. te Woerd, D. vd. Ploeg (1991). Methoden in het reken-wiskundeonderwijs, een rijke context voor vergelijkend onderzoek. Utrecht: OW&OC/ISOR.

Harskamp E. & C. Suhre (1986). Vergelijking van rekenmethoden in het basisonderwijs. Groningen: RION.

Hiele, P. M. van (1973). Begrip en Inzicht. Purmerend: Muusses.

Resnick, L. B. & W. W. Ford (1981). The Psychology of Mathematics Instruction. Hillsdale N.J.: LEA.

Thompson, A. G. (1984). The Relationships of Teachers' Conceptions of Mathematics and Mathematics Teaching to Instructional Practice. Educational Studies in Mathematics, 15 (2) 105-127.

Treffers, A. (1987). Three dimensions. A model of goal and theory description in Mathematics Instruction - the Wiskobas Project. Dordrecht: Reidel.

Wijnstra, J. M. (1988). Balans van het rekenonderwijs in de basisschool. Arnhem: CITO.

PROCESS–PRODUCT RELATIONSHIPS IN COMPOSITION: THE CASE OF REVISION

Uriël Schuurs, National Institute for Educational Measurement (Cito), Arnhem, & Huub van den Bergh, Faculty of Humanities, University of Utrecht, Utrecht, The Netherlands

In writing research a paradigmatic shift can be noticed from a focus on product characteristics to a focus on process characteristics of the writing process. Especially the work of Flower and Hayes (1981) underlines the importance of knowledge of the writing process. In the 1980's especially the subprocesses of planning and revision have been studied in depth. The typical study of these processes is carried out according a novice–expert paradigm (i.e. the distinction between skilled and less skilled writers; see for instance, Sommers, 1980; Hayes, 1985). Hence, the research question in these studies is: what are differences between novice and expert writers in terms of processes?
If we concentrate on revision it seems that there are quantitative and qualitative differences between novice and experts writers. Experts do not only revise more, but their revisions are also more directed at the meaning of the text. Novices on the other hand pay more attention to lower level revisions like spelling and punctuation. From this it goes without saying that the kind of revisions made by good writers have much more impact on text quality than merely the correction of mistakes. One of the aspects that make revision research interesting from a point of view of educational research, is the way it can improve the writing skills of novice writers (Hillocks 1986).

Revision is a complex process; it involves at least two subprocess. First a writer has to evaluate a text passage to some standard, extracted from his knowledge of the subject matter and the rhetorical problem situation. If a passage does not meet the writers standard the rewriting activity itself will be triggered. In this respect the Flower and Hayes model does not seem to be as sparse as can be: writing and rewriting activities do not differ in principle. Hence, there is no need to assume a rewrite process as subordinate to revision. Therefore, in the study of revision a distinction should be made between the evaluation and the correction of text passages.
Partly on basis of the argument above research results cannot be interpreted equivocal. First, many studies concern the correction of (errors in) text produced by the respondents themselves. However, different respondents are apt to make different mistakes, the interpretation of type and frequency of revisions is obstructed. Furthermore, the standards of writers to which they compare text

passages as well as their interpretation of the rhetorical situation differ, so that the actual revision processes in the Flower and Hayes sense differ. A last point worth mentioning concerns the distinction between extreme groups of writers (Cronbach, 1975). This might some problems in the interpretation of empirical studies.
In order to study differences in evaluation and correction of 'errors' it might be worthwhile to present the same 'errors' to respondents at different occasions. That way at least the problem situation and the type of 'error' are kept constant for all individuals. Furthermore, it might make sense to distinguish between three types of respondent: thosewho do not recognize errors, those who only recognize 'errors' but do not correct them, and those who recognize and correct 'errors'. Such a procedure allows for a distinction between evaluation (recognition of 'errors') and correction of 'errors' (rewriting).

Data
At two different occasions we confronted seventh graders with a text in which thirteen 'errors' or ambiguities were hidden (N = 162; Schuurs, 1990). We assume that there are differences between the scores on the first occasion and the second occasion three months later (as between both occasions an extensive writing course took place). More specifically, we assume that students who did not recognize errors at the first occasion (evaluation) will recognize and/or correct the errors. Students who only recognized errors at the first occasion will primarily be correctors afterwards, whereas no change is expected for those who corrected the errors from the start.

Analysis
The data can be described per item by means of a crosstable with two dimensions. On the first dimension the student is positioned at the start of the school year as one who did not recognize the 'error', as one who only recognized the error, or as one who recognized and did correct the error. At the second dimension, the student is positioned in the same way. So, per item we have a crosstable with two dimensions, with each three levels. Each crosstable was analyzed by means of a loglinear model (Fienberg, 1980; Van den Bergh, 1990).

Results
The data show in general that students tend to recognize and/or correct more errors at the second occasion as opposed to the first. So, generally speaking there is progression. However, we are meanly interested in the changes in type of corrections for the three subgroups (students who did not recognize errors; students who recognized errors but did not correct them; students who recognized and corrected the errors). Especially, the students who were not able to recognize errors at the first occasion showed large progression. This might be expected since only those students were able to improve 'two steps' (i.e. recognition and correction). Nevertheless for the majority of the items a significant proportion of these students only recognized the 'error' at the second occasion, and therefore improved 'one step'.
At the first measurement occasion about 12 percent of the students did only recognize the hidden errors and were not able to correct them. At the second measurement this percentage dropped to a mean of one percent, and only 1.1

percent did not recognize an error at the second occasion, but did so at the first measurement.
We take these results in favour for the plausibility to distinguish between recognition of errors (evaluation) as a different cognitive activity from correction of errors once they are recognized (rewriting).

References

Cronbach, L.J. (1975). Beyond the two disciplines of scientific psychology. American Psychologist, 2, 116–127.

Fienberg, S.E. (1980). The analysis of cross–classified categorical data. Cambridge: Mit press.

Flower, L.S. & Hayes, J.R. (1981). A cognitive process theory of writing. College Composition and Communication, 32, 365–387.

Hayes, J.R. (1985). Cognitive processes in revision. Eric documents (ED 267 396).

Hillocks, G. (1986). Research on written composition / New directions for teaching. NCTE, Urbana, Illinois.

Schuurs, U. (1990). Leren schrijven voor lezers (Learning to write for readers). WMW–3, Dissertation Un. Twente.

Sommers, N. (1980). The predrafting process of four high– and low–apprehensive writers. Research in the Teaching of English, 18, 378–388.

Van den Bergh, H. (1990). On the construct validity of multiple–choice items for reading comprehension. Applied Psychological Measurement, 14, 1–12.

HOW TO ACHIEVE THE GOALS OF ENVIRONMENTAL EDUCATION? CONSIDERATIONS FOR THE IMPLEMENTATION OF ENVIRONMENTAL EDUCATION AS A NEW CURRICULUM

J.W. Eberg, ISOR, Department of Educational Research, University of Utrecht, Utrecht, The Netherlands

From science education to environmental education: a new curriculum

In the Netherlands "natural history education" or science education dates back to the year 1857 when it became obligatory. It dealt with the principles of science studies in a verbalistic way and this situation remained unchanged until after W.W.II. Science education evolved into what we now know as environmental education under the influence of three environmental waves. This were, respectively, an ideological, a political, and an economical oriented wave. The latter is still going on. We have entered the age of sustainable development in which economical activities go hand in hand with ecological activities. Also in the last few decades the goals and activities of the "green" and (the younger) "grey" organizations for environmental education are more and more drawn together. Both the care of nature and the concern about environmental problems are part of the subject matter of environmental education. An important contribution to the creation of environmental education as a curriculum at elementary school level in the Netherlands was made by the NOB project (1978-1990) of the Institute of Curriculum Development (SLO). This project emphasized the importance of both coherences in the material reality and the learning and stimulation of knowledge, skills and attitudes. At the end of the project it was clear that science education alone is not enough to meet the social interest in nature and matters concerning the environment. Environmental education allows for an integral interest in nature and the environment, as well as the scientific and social aspects of these two areas. The newness of environmental education as a curriculum is that it can not be regarded a delimited subject or a well–defined type of course. It has interfaces with subjects as biology, history and geography, and with other educations such as cultural and developmental education, and health education. It forms a curriculum that is composed of the different environmentally oriented components of existing subjects and courses. This does not make the pursuit of the goals of environmental education any easier.

The goals of environmental education

According to the Institute of Curriculum Development (Pieters 1990) a main goal of environmental education at school is: to teach students how they can contribute to the advancement of a society based on sustainable development, starting from the interrelation between man and the environment. For primary education five more specific teaching goals can be derived from this. These goals are concerned with the significances of nature and the environment for man, the influence of man on nature and the environment, environmental problems, solutions to environmental problems, and careful association with nature and the environment. The goals are meant to be achieved simultaneously in a way that learning experiences become more and more extent during school time. These teaching goals give scope for a variety of teaching activities and a wide range of subjects. The teaching activities may also enclose cooperative learning, working with lesson boxes, drama, and fieldwork. Research that has been carried out recently (Eberg, Eijkelhof, Kortland, Stokking 1991) shows that working with organic material as well as discussion and discourse are characterized as highly usable means of teaching environmental education. In order to make things easier to understand for children and to maintain the coherence between school subjects one could start with models such as Self–Society–Surroundings (S–S–S) and Head–Heart–Hands (H–H–H). The latter stands for acquiring knowledge, developing attitudes, and practicing insights. It also stands for the main three aspects of learning goals, i.e. cognitive, affective and behavioural development.

Concept development as a cognitive goal

Smith & Medin (1981) divided theories of concept learning into classical, probabilistic and exemplary theories. However, there is no ideal theory of concept learning. Concepts and concept learning are more complex than we tend to believe. Many researchers have been instrumental in the creation of useful theories on the subject. The most important amongst these are the theories of Piaget, Ausubel, Bruner, Vygotsky, Davydov, and Keil. The acquisition and development of concepts varies from person to person. The notions and beliefs of children and the way they handle concepts are particularly difficult to structure. Children's ideas are often labelled as misconceptions or naive theories, which have to be changed by school science. Keil shows that for children the differences between natural kinds, nominal kinds and artifacts seem to be more a matter of dimensional continua then of the presence or absence of specific features. Keil also argues that there is a framework of conceptual categories and causal beliefs that enables even very young children to understand kinds at a deeper, theoretically guided level. Davydov, too, argues in favour of the development of theoretical thinking in primary education. In an ISOR study of concept development in environmental education at primary schools (Boekenoogen, Eberg, Stokking 1991) the authors came to the conclusion that children use and control the selected concepts fairly well. This was often the case even before the lessons on the concepts were given. It also appeared that a literal formulation of questions and avoidance of obscurity about the subject strongly affects the answers of the children, particularly of the younger ones. On the whole children have a passable environmental awareness. The researchers are under the impression that the children were treated cautiously and that they are capable of handling more complex subjects. For that the type of

education should be developmental in stead of following. Also the didactical skills of the teacher play an important role in the process of concept development. Didactical implications for concept development are: working with models, stimuli for concept learning, 'realistic education', cooperative learning, discussion and discourse, learning supported by illustrations in visually rich contexts, and gaming.

Attitude development as an affective goal

An attitude is an learned predisposition to act in a consistently favourable or unfavourable manner with respect to a given object (Fishbein & Ajzen 1975). Attitudes are not innate but acquired and they can change. They can be influenced by a number of relational factors (Stokking, Boekenoogen, Eberg 1991). Social learning theories and cognitivist theories speak of moral education. Kohlberg (1969, 1973) presents a theory on the development of moral values and norms of children which makes it clear that moral education is an active process which can be enforced by schooling. Several value–forming strategies are distinguished, in particular strategies that are of interest for environmental education (Caduto 1983, Delhaas & Koekkoek 1990, Buisman 1991). For a value–forming strategy to be effective in developing attitudes the results must be measured. However, value–transfer is difficult to measure. Problems that occur in measuring values and attitudes are related to distinguishability, stability, and generality. Consequently, it is advisable to use different means for measuring attitudes. De Groot (1986) recommends the learner report. Another means of measuring attitudes is the use of scales. In an ISOR pilot study concerning the evaluation of environmental education (Stokking et al 1991) the researchers found that in studies on the measurement of attitudes towards environmental subjects there is little systematic towards the subjects that should be measured and the learning effects involved. Also the measurement instruments found in literature are not optimal from a technical point of view. What must be done is educational research into both the development of suitable and valid instruments to measure attitudes towards environmental subjects, in particular aimed at children, and into the effects of values forming strategies as a part of environmental education. For teachers who want to be active in applying values forming or attitude development, didactical implications are: handling moral dilemma's, using a didactical model for the improvement of experiences from nature, essay writing, discussions, reasoning and debating, drama, and gaming.

Environmental care at school as a behavioural goal

Environmental care at school may be described as the sum–total of school activities that are environmentally sound. It is generally regarded as a matter of "doing things", which demands motivation and alertness, coordinated organization of responsibilities, and a school environment plan. The plan consists of a variety of ecological activities to be carried out within a certain period. There would seem to be several problems connected with environmental care at school, given that sloppy thinking and carelessness are not easily changed. One could think of trying to influence behaviour. From theories of social psychology (Ajzen & Fishbein 1980, McGuire 1985, Rogers 1983) several ways of doing this could be extracted. However, for environmental care at school (and for environmentally sound behaviour in general) to succeed, concrete perspectives for action must be offered.

When pupils see direct results they maintain their motivation. During the ISOR study of concept development in primary schools (Boekenoogen et al 1992) it appeared that many environmentally sound activities are already self–evident to the younger generation. Perhaps these processes indicate that in the same way that health education has resulted in sanatary conditions, environmental education will result in sustainable development. Also policy strategies for environment directed education by the school management can contribute to environmental care at school (E. Marx 1990). Didactical implications in relation to environmental care at school are: decision–making skills, interviewing, fieldwork, lobbying and exerting influence, and results directed action.

Future research and development in environmental education

Research in the field of environmental education is limited but growing in the Netherlands. Significant efforts are being made in the following fields: educational equipment (Van den Akker, Reints), experience of nature (Margadant), curriculum development (Boersma, Pieters), fieldwork (Schermer), diffusion and implementation (Leussink, Schermer), value–formation (Delhaas & Koekkoek), and cognitive development (Van Oers). Additional research will be required to support the general implementation of environmental education as a new curriculum. Specific support is needed for the development of teaching methods in environmental education. In this connection Eberg et al (1991) carried out a study on programming problem areas which require attention in the next four years. Their study resulted in research proposals for primary and secondary education. Experts that were questioned in this study stated that not only the implementation and didactics of environmental education in the schools need to be supported. Moreover the lack of interest, knowledge, skills, and experiences of teachers in regard to environmental education calls for refresher courses, and proper programs and didactics in teacher training. Besides educational research, curriculum development is of great importance. Cooperation between research and development for the benefit of environmental education is advisable. The concerted work of researchers and developers could be described as developmental research. It is not inconceivable that also school attendants and teachers join this multi–disciplinary group. The main goal of the combined efforts of all these educational workers involved is an effective implementation of environmental education and that, as a result of that, this new curriculum will contribute to a sustainable society.

References

Boekenoogen, M.S., Eberg, J.W. & Stokking, K.M. (1992). Theorie en praktijk van begripsontwikkeling bij natuur- en millieu-educatie in het basisonderwijs. Enschede: SLO.

Eberg, J.W., Eijkelhof, H.C.M., Kortland, J. & Stokking, K.M. (1991). Naar een didactiek voor natuur- en millieu-educatie. Utrecht: ISOR/CD-B.

Eberg, J.W.& Stokking, K.M. (1992). Inplementatie en didactiek van NME in het basisonderwijs. Een methode voor de ontwikkeling naar waarden. Onderzoeksplan. Utrecht: ISOR/HMN.

Stokking, K.M. , Boekenoogen, & Eberg, J.W. (1991). Evaluatiemethoden voor natuur- en millieu-educatie. Een vooronderzoek. Utrecht: ISOR.

127

BEST COPY AVAILABLE

EXPERIENCES IN ENVIRONMENTAL EDUCATION CURRICULUM RESEARCH AND DEVELOPMENT IN THE NETHERLANDS

M. Pieters, National Institute for Curriculum Development, Enschede, Netherlands

In the 1970's and 1980's, environmental education (EE) was a marginal activity in formal education in the Netherlands. Some schools did extracurricular activities, such as project work. Nongovernmental organizations for environmental protection produced teaching materials as a means of putting environmental education on the teachers' agenda's. However, it remained a marginal part of education, and from dissatisfaction with this situation NGO's looked for cooperation with institutes experienced in developing curricula. With support of several ministries a cooperation started in 1986 between some NGO's, the University of Utrecht (Physics Education Group, presently part of the Centre for Science and Mathematics Education) and the National Institute for Curriculum Development (SLO).

After a period of developing, revising and publishing teaching units in order to quickly offer some good examples of what could be possible, a more research oriented strategy was developed. This strategy started by making an outline of a core curriculum for environmental education in primary and secondary education, based on theories from environmental and from educational research. Then the conceptual framework and skills proposed in this core curriculum were used in identifying 'niches' for environmental education within curricula of regular school subjects. Partial curricula were written for environmental education in several school subjects. In 1990/1991, parts of these were elaborated into experimental teaching units for secondary education. The subjects involved are geography, biology, and physics/chemistry in lower secondary school, and biology and physics at upper secondary level. In the first round of research, in 1991, each unit has been tried out by two teachers.

The research aims at answering the following questions:

- What activities contribute to a better understanding of the concept of sustainable development?
- What activities contribute to improving a thoughtful decision-making by pupils on environmental issues?
- How to take into account present ways of reasoning of pupils about environmental issues?
- What new professional skills do these activities demand from teachers?

Answers to these questions will be used in revising the EE curriculum designs, in designing and contributing to teacher training activities, and as contributions to a more general didactical theory of environmental education.

The first round of research has been carried out by:

- interviewing pupils in order to get insight into common preconceptions and ways o reasoning in a specific environmental issue;
- designing a teaching unit on the basis of these interviews, of the subject's curriculum, of teachers' suggestions and of the EE curriculum proposal; the designs were based on a constructivist view of learning processes;
- agreeing with the teachers upon crucial activities and their planning;
- collecting data on pupils' preconceptions and ways of reasoning by means of a questionnaire for all pupils (pretest);
- observing in classrooms during execution of the selected activities;
- collecting data on pupils' preconceptions and ways of reasoning by means of a questionnaire (posttest);
- interviewing the teachers on their experiences and evaluation of the lessons.

The first round of research was not entirely satisfactory. The questionnaire turned out to be too rough to give insight into ways of reasoning beyond the answers. Some of the designed activities did not succeed to keep all pupils' attention. And the teachers' interpretations of the contents of the unit and the research questions departed from what the researchers had in mind.

These complaints gave rise to modifications in the units, especially by giving mor emphasis to pupils' preconceptions. Also, format and contents of questionnaires were adapted, using statements made by pupils in the first round.

In June 1992, fresh impressions can be given from the second round of research. We intend to focus on two discussion topics:

- what possibilities and what restrictions can be seen in this research based strategy for curriculum development?
- what methodological problems are met in this kind of research?

MAN AND THE ENVIRONMENT: THE ONGOING CURRICULUM DEVELOPMENT FOR PUPILS 15-17 IN GENERAL EDUCATION (1991-1994)

Margherita Landucci, Istituto Tecnico Industriale Statale, Pordenone; Hugo E. Lombardini, Centro Linguistico Interfacoltà Università degli Studi, Venezia, Italy; Marc Renyier, Lycée Polivalent Jean-Moulin, Albertville, France, & Fernando F. Rojero, Istituto de Bacillerator Las Llamas, Santander, Spain

Teaching requires communicating. Not all teachers possess the communication skills necessary to teach. Using a communicative context we are developing a multidisciplinary curriculum which stems from interactions between pupils and their social, cultural economical, political and natural environment. Thanks to the European Commision Communities - who started and are now assisting with their financial support pilot projects on international class exchanges - we are experimenting with and suggesting for use use in public schools teaching methodologies which are not yet a part of the scolastic curriculum: environmental education and chemical linguistics.

The issuer that schools traditionally make available to transmit knowledge is the teacher and the textbook, and in more fortunate cases, the video. In our proposal the issuer that transmits knowledge is the same environment in which man lives and in which the pupil can have direct experiences through guided visits. The role of the teacher is that of interpreting that particular environment in multidisciplinary scientific ways, thus realizing the similarity between the demands of the curriculum and the evidence from empiric sources. A suitable choice of environmental typologies allows:

- emphasizing the role of the environment seen in the context of the environment in general: *environmental education;*
- covering the contents of a regular scolastic curriculum: *education through the study of the environment.*

From here arises the necessity to find new methodological routes. That which we suggest singles out linguistics - and the science of language in general - in a series of elements that we claim valid and that we propose to develop, not only in this work that we consider as a presentation of objectives, but also in successive works. It it is true that linguistics imposes schemes and values on reality, the knowledge of the way in which the language functions in the aspects considered by different disciplines must be utilized to the advantage of the pupil. The

development of scientific language in the pupils must not be produced by creating a completely new structure, but instead by emphasizing the structures already existing in specialized fields, or structures rarely used which could be developed in programmed ways. In the teaching field in the disciplines of science and chemistry in particular, we find four problematic areas which are well delimited and difficult to solve:

- The communicative situation established between teaching sources and the pupil.
- The clarification of the close relationship that exists between reality and the science of chemistry, in order to overcome the concept the students may have that chemistry consists only of numbers, formula and calculations.
- The interpretation and the use of chemical codes in respect to the rules within the code.
- The passive reflection *(comprehension)* and the active reflection *(internalization)* that indicates inconsistencies of the system (due almost always to a lack of information).

The prospective that today we propose to surpass the difficulties mentioned above are as follows:

- The utilization of concepts from the linguistic and literary disciplines for a better comprehension of the pairing of *Expression-Content* in chemistry.
- The utilization of the active knowledge which the pupil has in his mother tongue to the benefit of his comprehension of the language of chemistry.
- The utilization of various concepts that the pupil has used in linguistic or in literary disciplines.
- The development of pupil's capacity to recognize icons and symbols, for example those in a card game, that are already part of his cultural background.

From this prospective we can develop, among others, the following concepts:

The message

To be functional, reference must be made to a context that can be comprehended by the receiver. This requires a code that must be, at least in part, common between the issuer and the receiver, and a contact - a physical channel and a psychological connection that activates and maintains the communication. The duty of the teacher is to control the *distance* between the codes that were used in the codification and the codes which the pupil possesses, in order to minimize the *noise* effect in the transmission of the content.

The encyclopedia of the pupil

that is used as a starting point for the first approach to the scientific and technological world. With this first approach the pupil must acquire some textual competence which refers to the reality considered. Thus when the pupils approaches other detailed areas, they will use their enriched encyclopedia as a base to develop other textual abilities which, when combined, will produce the general textual competence of science.

The context
In order that, for example, the iron ores *magnitite* Fe_3O_4, *ematite* Fe_2O_3 and *siderite* $FeCO_3$ transform into molten iron they must be placed in a suitable context. The creator of this context is the blast furnace. Six semantic categories describes the possible roles carried out by the characters in a tale. The *assistant* aids the *subject* in the trials that must be surpassed to attain the coveted *object* and in which the obstacle is the *opposer*. The *issuer* places the object as the aim of desire and the purpose of communication, while the the *receiver is* the one who benefits from it. The scheme is available for the most varied semantic investments. In the blast furnace the iron ore (Subject) is made to surpass defined trials to become iron (Object). Oxygen (Oppositor) is won over thanks to the intervention of Carbon (Assistant) and eventually the protagonist performs its final duty due to a need (Issuer) and in favour of man (Receiver).

The tale and the plot
It is noted that the tale (fable) taken in a linguistic sense is accessible to everyone because it uses a type of temporal approach and the receiver has the temporal code. On the contrary the plot is known only by those who also have the textual competence. Thus we can not present a chemical or technological argument with a different plot from that which adapts perfectly to the tale, otherwise the only thing that we obtain is a loss of information due to 'noise' caused by not knowing the code. The physical reality can be achieved from a variety of point of view. This means that the same story is told with different plots. The same blast furnace may have many different plots according to the purpose of teaching that the issuer pursues. For this it is useful to present a story which has a plot that corresponds perfectly to the tale, seeking to use ellipses (the way in which the plot plays with time) that will later serve to develop the subplots. These subplots will permit the parallel development of the true scientific plot, the knowledge of particular codes and the consequent application of the textual competence. When a general story, GS, is used together with a variety of ellipsis there is the possibility a) to return to the general tale every time it is necessary, and b) to interrupt the development of an ellipsis to acquire definite fundamental information from another ellipsis.

GS → ellipsis2 → GS → ellipsis7 → GS → ellipsis3 → GS → ellipsis1 → GS → ellipsis4 → etc.

This route permits the knowledge of phenomena not only in their complete form but also in their systemic organization.

Semanteme, lexeme, moneme, morpheme, grapheme, etc.
In the chemical reaction

$$CO(g) + Fe_3O_4(s) \rightarrow CO_2(g) + 3FeO(s)$$

The formula *3FeO(s)* that expresses the iron oxide called wurstite contains six graphemes *3, Fe, O, s,* (,), a morpheme 3, a composed lexeme *FeO(s)* and three monemes *Fe, O, s*. We speak of a composed lexeme because it could also be understood as two lexemes, i.e. the lexeme *Fe* and the lexeme *O*. In chemical language a deliberate mark is used to distinguish different lexemes composed of

equal simple components. This mark is analogous to what is noted in English as a hyphen, for example in 'social-economic'. These marks are the subscripts. The difference is apparent between the compound *FeO* and the compound Fe_3O_4. Considering now the problem of the morphemes, in the formula *3FeO(s)* the *3 is* present only because it is found inside an equation where the quantity by definition must remain equal in both members, as by convention the principal verb of a sentence must agree with the subject.
In approaching the problem of semantemes, we can extract from the lexeme *FeO(s)* the following semantemes: the relative molecular mass or molecular weight $M_r = 71.859\ mol^{-1}$; $t_{fus} = 1369°C$; the crystal lattice *CUB*; the density $\rho = 5.7 g\ cm^{-3}$; etc.

All this indicates how the utilization of concepts from the science of linguistic makes the initial comprehension of the argument easier, permits a deeper level of learning, announces tha unity of scientific knowledge and of human science, and gives incentive to the advent of interdisciplinary work which we have as much hoped as abandoned.

References

A.A.VV., (1986). Holman, J. (Ed.)Science and Technology in Society. The Association for Science Education, UK.

AA.VV. Landucci, M. Rendina N. a cura di, (1988). Scienza e Tecnologia nella Società. Satis, Pordenone.

Barthes R., Greimas A. J., Eco U. et al., (1985). Análisis estructural del relato. Premia, Puebla,.

Benveniste E., (1985). Problemas de linguistica general I-II. Siglo XXI editores, Mexico.

Dorin H.,Demmin P. et al., (1991). Chimica. Lo studio della materia. Arnoldo Mondadori, Torino.

Ducrot O., Todorov T., (1986). Diccionario enciclopédico de las ciecias del lenguaje. Siglo XXI editores, Mexico

Eco U., (1985). Tratado de semiótica general. Lumen, Barcelona.

Genette G., (1986). Figure III. Discorso del racconto. Einaudi, Torino.

Halliday M.A.K., (1978). Language as social semiotic. Edward Arnold, London.

Kerbrat-Orecchioni C., (1986). La enunciación. De las subjetividad en el lenguaje. Hachette, Buenos Aires.

Landucci M. & Lombardini H.E., Una nuova prospettiva per lo studio delle scienze. [a proposal].

Landucci M. & Poles P.G., (1991). "Dimitry": a computer card game. Atti del I Convegno Nazionale di Informatica Chimica, Venezia 1-3 Marzo.

Lombardini H.E., La noción de función en la ciencia del lenguaje, in Letras.

Marchese A.,(1989). Dizionario di retorica e di stilistica. Arnoldo Mondadori, Milano.

Segre C.M, (1985). Avviamento all'analisi del testo letterario. Einaudi, Torino.

A REVIEW OF THE ADEQUACY OF CURRENT STRUCTURAL AND LOGISTIC FACILITIES FOR THE IMPLEMENTATION OF ENVIRONMENTAL EDUCATION IN A SAMPLE OF EUROPEAN COUNTRIES

Walter D.S. Leal Filho, Institute of International Comparative Education, University of Hamburg, Hamburg, Germany

Abstract

A recent recommendation of the Committee of Ministers of the Council of Europe proposes that the governments of member States undertake a number of structural changes and carry on various curricular improvements, in order to cater for the implementation of environmental education at their school systems. The likelihood of the implementation of the items of the recommendation among members countries, however, it is not fully known.

Based on the need for research, this paper presents an analysis of the structural facilities currently available for the fostering and/or initial implementation of environmental education in the teaching system of a sample of six European countries (Great Britain, Germany, France, Portugal, Greece and Hungary). It also critically looks at the requirements needed to meet the Committee of Minister's recommendation and the extent to which individual countries are able to meet them. Suggestions as to how various items from the recommendation may be made operational, are made.

Introduction

Environmental education is becoming increasingly popular in Europe, its growth being happening partially as a result of the concern showed towards environmental conservation and maintenance of quality of life in the continent (Leal Filho, 1992). In a demonstration of recognition of the need for environmental education and of awareness of its relevance to the environmental conservation process, the Committee of Ministers of the Council of Europe recently approved a recomendation aimed at fostering environmental education among the countries members. The recommendation, which was issued based on Article 15 b of the Statute of the Council of Europe, has a great relevance in the sense that it suggests ways through which environmental education may be wide spread in Europe. Among other items, the Recommendation states that governments of the States members of the Council of Europe should bear in mind the "basic principles for the promotion of environmental education" gathered from the various events

(Tblisi, Moscow, etc) which shaped up this methodology, when drawing up or reviewing their environmental education policies; ensure that teachers have an opportunity of revitalising and diversifying classrooms activities and basic learning processes by providing them with (teaching) resources and promote working relations with research institutes in the field of natural science, social science and education.
But in order to verify whether the Council of Minister's recommendations are likely to be fully implemented, research works on the extent to which structural and logistic facilities are found in individual European nations are necessary. They are also timely procedures in order to identify, current trends, and to guide eventual improvements that may be needed in fulfilling the recommendations.

Methods
Based on the need for research, an empirical investigation aimed at identifying the extent to which key items related to the execution of environmental education programmes are found in European countries was undertaken. The study involved a sample of five distinct nations, namely Germany, England, France, Portugal, Greece and Hungary. They were selected on the basis of their different levels of evolution in terms of environmental education and on the access granted by educational authorities in the selected countries, to relevant information. A checklist consisting of eight criteria was employed to assess the existence of various structural and logistic facilities in each nation. They are: the status of environmental education in the formal curriculum; the existence of regularly-run in-service teachers training programmes; the existence of teaching materials (e.g. books, booklets) especialy prepared for the environmental education of schoolchildren; the existence of teaching resources for environmental education (e.g. audio-visuals, posters, charts, videos, films); the provision of financial means for the execution of environmental education works; the existence of facilities for field studies (field studies centres, outdoor centres, etc); the presence of advisors for environmental education and, finally, the existence of monitoring strategies through which on-going works could be followed and through which eventual problems could be detected at an earlier stage.
Data from individual countries were collected by means of analysis of governmental reports, academic theses and visits to schools, field studies centres and research institutions in various parts of the selected nations. These were complemented by a collection of bibliographical records on advancements in environmental education from each area. Talks to experts from the selected countries were also used as data gathering procedures, having the advantage of bringing first-hand statements from people who are in continuos contact with a country's context and are, as a result, well prepared to provide information on current trends.

Results
Table 1 shows the various differences in terms of the availability of structural and logistical facilities are seen among the countries surveyed. The formal existence of environmental education, for example, is not seen in Portugal or Greece, confirming Episcopopoulou (1991). It is now being developed in Hungary which still suffers significant problems (Jakucs & Lakatos 1990) and is present in the

other nations studied. In addition, teachers' training programmes, which are present in the first three nations, are yet to be made widely available in Hungary and Greece, whilst Portugal simply does not have them.

Table1
Distribution of structural and logistic facilities for environmental education in the sampled countries

Item	Germany	England	France	Hungary	Greece	Portugal
Existence in the curricula	present	present	present	being developed	absent	being developed
Teachers' training	present	present	present	restricted	restricted	absent
Teaching materials	present	present	present	restricted	restricted	present
Advisers for EE	present	present	present	absent	absent	absent
Resource	present	present	present	absent	restricted	restricted
Finances for EE	present	present	present	present	absent	absent
Field studies centres	present	present	present	present	absent	absent
Monitoring strategies	absent	present	absent	absent	absent	absent

Both findings are very significant in curricular terms, as the provision for the introduction of environmental education initiatives in formal teaching is a pre-requesit for its implementation.
Without this, it becomes very difficult to systematically conduct teachers' training schemes focussing on environmental education and conservation. it is also seen in Table 1 that the existence of teaching materials is very restricted in Hungary and Greece, whilst in Portugal these materials are surprinsingly found.
A further evidence of the unequality seen among the countries investigated is that environmental education advisers and resources are concentrated in the first three countries, being absent or restricted elsewhere. Furthermore, albeit present in Hungary, field studies centres havenot still been implemented in Portugal or Greece. In terms of monitoring strategies, England seems to be the only country to

have made formal provisions for them, despite their relevance in assessing curricular developments and progresses.

Discussion and conclusions

Data gathered through the study indicates that there various inequalities among the sampled countries. Those illustrate the degree of variability which exist in terms of the emphasis givento environmental education. It is seen that the value of the Council of Minister's Recommendation may be increased even more if they are complemented by Europeanwide working strategies aimed at reversing the inequalities and to a considerable extent, the deficiences, here identified. To reach such a goal and to fully meet the Council's aspirations, it is relevant that a number of procedures be adopted. They include: the establishment of Europeanwide initiatives of implementation of environmental education, possibly coordinated by a central supra-governmental body. An European Council for Environmental Education could be set for this purpose; stronger efforts from individual countries towards the nationwide implementation of environmental education, including from the ones with economic difficulties, which should make proper budgetary and curricular provisions for the introduction of environmental education in their teaching systems; the establishent and, as applicable, improvements, on in-service and pre-service teachers training programmes, special recycling courses for those teachers who cannot take part in normal training schemes; provisions for environmental education materials and resources in each country.

Despite their structural differences, England, Germany and France offer some examples of strategies that may be successfully implemented and that may be used as a basis for developments in the countries where environmental education is still in an embrionic stage.

References

Episcopopoulou, C. (1991) Environmental education as an educational experiment in a Greek school. Paper presented at the "International Conference on Environmental Education". DES. Manchester, March, 1991.

Jakucs, P. & Lakatos, G. (1990) The ecological aspect of environmental education in Hungary. Higher Education in Europe, 15 (4), pp. 24-28.

Leal Filho, W. D. S. (1992) Environmental education in Europe. European Environment. IN PRESS.

CURRICULUM AND EUROPEAN DIMENSION: WHAT CAN PUPILS IN CZECHOSLOVAKIA LEARN ABOUT THE NETHERLANDS

J. Prucha, Institute of Educational and Psychological Research, Charles University, Czechoslovakia

Introduction

The first part of this paper explains profound changes in curricular policy in Czechoslovakia. The changes have been caused by the systemic transition of the whole educational system following the "velvet" revolution in Czechoslovakia, November 1989. With respect to curricula the most important changes concern:

- Teaching of Marxist doctrine was abolished in all levels of schools. Ideological subjects and topics that were a part of the previous curricula have been replaced by new subjects including social sciences, survey of world cultures and religions, ethics, democracy and humanity~ family education, etc..
- Textbooks on history, civics and some other subjects which were used before November 1989 have been removed from schools. This creates the necessity to develop new curricula and textbooks for the respective subjects.
- Teaching foreign languages has been radically changed: Until 1989 Russian was the first and for all students an obligatory foreign language in all types of schools. Instead, students can now freely choose among English, German, French and some other languages (see, in more detail, in Prucha, 1991).

A new curricular policy in Czechoslovakia is resting upon new educational legislation. Legislative measures in education have introduced some important changes, in particular:

- Pluralism in the school system. Under the totalitarian system all schools were state run. No alternative schools were allowed. At present, a pluralistic arrangement is realized step by step, i.e., alternative schools (private schools and schools run by churches) are recognized as legal counterparts to state schools.
- Decentralization of education. According to new legislative acts schools possess rights of self-management. This concerns, in particular, universities which became autonomous and independent of the central management. School-based management means that teachers, parents and local authorities have a great decisive role in realizing curricula.

One of the important changes in curricular policy in Czechoslovakia concerns the "European dimension" in education.

The second part of the paper deals with problems of introducing the European dimension into curricula. The European dimension is understood in the sense of

proposals for a core curriculum suggested by Hooghoff (1990). In this paper, the European dimension is explained by means of two empirical studies.

Aim

The aim of the investigation is to explain :

- How the topic "Europe" (using the Netherlands as the case) is included in Czech textbooks?
- How this topic is reflected in pupils' knowledge and attitudes. The investigation has two phases applying different methods?

Method

A sample of 12 Czech textbooks on history, geography, and civics of the basic (comprehensive) school, grades 5-8, have been analyzed. The method applied was content-analysis. All information related to The Netherlands (history, economy, culture, etc.) was excerpted and analyzed.

A questionnaire was submitted to a sample of Czech pupils of grades 5-8 (age 11-14 years). The sample covered four schools representing typical educational areas (the Prague city and rural communities). The questionnaire involved several parts:

- Questions inspecting pupils' knowledge about the topic "The Netherlands" (political, social,economic aspects).
- Questions reflecting cultural aspects of "The Netherlands" (people, language, religion and ethical relations).
- A non-verbal task (drawing) by which pupils indicated the geographical position of Czechoslovakia and The Netherlands in Europe.

Results

At present, only preliminary findings from textbook analysis are available. The results are not very satisfying. With the exception of the textbooks on Geography, all other textbooks of the given sample do not present any important information about The Netherlands. In some textbooks on History there is, in several cases, only fragmentary information about The Netherlands (e.g., concerning the historical role of The Netherlands as a colonial super-power).

The findings of the second empirical study (questionnaire) will be presented in the paper.

Conclusion

A preliminary conclusion is that the relevant textbooks used, at present, in Czech comprehensive schools do not provide pupils with a knowledge about The Netherlands. However, the same is valid of other minor European countries as, e.g., Norway on Finland. Thus, it can be stated that the official curricula of the Czech comprehensive schools do not satisfy demands for introducing the "European dimension" in education. However, this finding should be compared with data reflecting pupils' real knowledge about the respective topic. A comparison will be presented in the paper.

References

Hooghoff, H. (1990). The Topic 'Europe': A proposal for a core curriculum. Working Paper A. National Institute for Curriculum Development (SLO): Enschede.

Prucha,J. (1991). Trends in Czechoslovak Education. Country Report. presented at the Workshop Recent Trends in Eastern European Education. DIIPF, Frankfurt a.M., June 5-7, 1991 (in print).

PROBLEMS OF CZECH CURRICULUM DEVELOPMENT: THE FOREIGN INSPIRATION IN THE NEW CZECH TEXTBOOKS

V. Martinkova, Institute for Educational and Psychological Research, Charles University, Czechoslovakia

This paper deals first of all with the relevant questions of Czech curriculum as they have crystallized since November 1989.
First, it shortly summarizes the present situation in the Czech educational system, and then it focuses on planning, designing and introducing the curriculum. With regards to the fact that "curriculum" is for the Czech pedagogy quite a new term, this paper shows also the difficulties which may appear and points also to the possible and necessary ways for further development.
In the following parts attention is concentrated on a very important problem of the curriculum - on textbooks. This paper includes short characterisations of the present state of Czech textbooks and the expected trends of development.
The part dealing with the foreign inspiration of Czech textbooks is considered to be the most important one. Although we are aware of the fact that the tradition of Czech textbooks is rich and valuable, at present we are facing a task to create new and good Czech textbooks. We look to highly developed countries not to "copy" them, but rather to take inspiration from them.
For this paper I have conducted comparison of Czech textbooks with chosen German, Austrian, English, French, American and Canadian textbooks, and I have come to the following conclusions:

- First of all it is not possible to discuss the textual aids in the form they were discussed in the Czech milieu. We must talk about audio-aids and (audio-)visual aids because the text books are usually completed with tape-recordings, audio recordings, computer programs, textual and pictorial folios for over-head projectors, with materials for creative activity etc..
- At the same time it is no longer sensible to categorize the school textbooks as textbooks in the narrow sense of this word, practice books, manuals, etc., because the character of the textbooks is usually quite varied and particular types of teaching materials very often go along with the text, which has got didactic and methodological character.
- This is connected with the fact that textbooks don't originate as independent books, but as sets of series of structures, which are rich in synchronic, for one form, and also diachronic level, for a number of following up classes.
- The former preferred cognitive approach loses its leading position in favour of the creative approach and know-how (practical knowledge, processes, methods, skills).

- Textbooks represent an integral part of curriculum and when solving their character it will be necessary to respect the curriculum as a whole.
- There are some ways of further development of the communicative aspect in Czech textbooks.
- Also printing industry, which will be, we hope, at least partly modernized in the near future will be able to satisfy the requirements of the graphical and typographical level of modern textbooks.
- There is another term which appears in the Czech educational milieu - it is the plurality (alternativity) of the textbooks. There used to be only one textbook for each subject in each form, sometimes even for different types of schools. Today their number is starting to rise thanks to the Ministry. In connection with this this is the first time that the textbooks don't represent to compulsory thinking which should be passed on to students, but instead they should offer various possibilities for schools, teachers and pupils.
- There is still a difference between foreign textbooks and Czech textbooks as far as the authorship is concerned. While foreign textbooks are usually written by large teams of experts and each of them is responsible for a certain part of the book or a defined task, in our country the textbook (or even a series of textbooks) is usually formed by one or two authors or a small team, which does not have to follow one conception. It is probably caused by lack of prestige of this creative activity, it is also related to the low financial compensation of the authors.
- The professional aspects of the textbooks are in our country understood very strictly as a relation of the material in the textbooks to the present scientific knowledge in the branch (subject). An integral part of other branches, especially of pedagogy, psychology, sociology, linguistics, printing, arts, ect. is not involved.
- A great perfection has been achieved in the textbooks of western countries in the sphere of logic and structure, where not only the basic means of building of the text are being used, but various means of graphics in the broadest sense of the word. At this point Czech textbooks are only in their infancy.
- We should see changes in approaches, from the transmissive approach (only relaying facts and knowledge) to the communicative approach, which focuses on the activity of the pupils and of the whole educational process. We should be aware of the importance of fun and play in foreign books, not seen in our textbooks.

I made use of primary and secondary materials from the sphere of textbooks and other materials of the Czech decisive sphere.

The research was carried our in two corresponding levels - in the level of theory and in the level of practice. In the theoretical part I focused my attention on the analysis of certain methodological materials and textbooks of humanistic subjects from various countries and on their comparatistic evaluation. I expressly formulated the main principles. In the practical part I conducted the pedagogical research, which included the method of observation, the method of experimental balance, discussions, questionnaires and pedagogical experiments. I focused on the function of textbooks and their aspects in the educational process.

A few final words: I believe that the Czech school system will return within a few decades to its position among Europe's educationally advanced countries, where it always used to be, and that it will continue in the noteworthy Czech pedagogical tradition. I would like to cite a passage from Comenius's Didactics: "To teach

means not simply to fill students with words, phrases, sentences, but to open their minds so that fruits, flowers, leaves, twigs, branches of knowledge may flourish..."

REFRAMING THE CURRICULUM STRATEGIES IN POST-REVOLUTIONARY ROMANIA

Anca Dumitrescu, Institute for Educational Sciences, Bucharest-Romania

In Romania, the Education Reform is not an independent project; it is a component of a broad, complex program for the Romanian society's deep-going transformation.
The challenge of educational change in Romania is twofold: it needs to maintain and strengthen what is good and valuable in the long traditions of our school-system, but also to stimulate and foster what is new and ever more required by today's and tomorrow's society.

During the transition period we are living in, the most important changes worked out in the education field have taken place at the level of school curricula. Among them we shall discuss the following ones:

- taking all the old ideology elements out of the curricula, so that they will no longer express the exclusive viewpoint of an unique totalitarian party;
- reducing the education contents by means of simplification and their more rational organization;
- using mother tongues by the existing minorities at all instruction levels and introducing the study of certain peculiar aspects regarding the history and culture of these minorities;
- introducing optional subject-matters from the 8th grade of middle school;
- assuring a minimum of general training for all the branches of post-primary education by a common core of subject-matters, so that in the Science section, as well as in the Humanities section 5 large groups of school-subjects are taught: humanistic education, basic scientifical education, physical training, educative courses and optional courses. The lesson number for each of these groups varies according to the specific targets of the respective secondary school education.

But after these initial changes of adjusting and adapting the existing curricula to the new conditions of post-revolutionary Romania, greater efforts should be made for elaborating new school curricula, which represent an essential stage in the Romanian Education Reform.
The analysis of today's school curricula for primary and secondary education realistically compared to various curricula existing in other countries, especially the European ones, has pointed out certain features which are worth taking into account in reframing the Romanian curricula strategies: they are rather exhaustive in information, but lacking the interest in developing the pupils'skills to use it

creatively and freely, they are strong on disciplinary detail, but weak on issues which span disciplinary boundaries, out-dated in presentation and missing many of the subjects necessary for a dynamic market economy and an open participatory democracy. Negative features such as: too prescriptive, undifferentiated, overloaded, inward-looking, etc. will be analyzed in our paper, too.

The main goals of curriculum change and development in Romania should follow 2 directions: improving the content of the already existent subject-matters and introducing new ones.

The first approach will focus on :

- renewing and broadning curricular objectives in close connection with the demands of present-day and future democratic society of Romania;
- incorporating new concepts and outlooks, scientifical discoveries, statistical and computing techniques within the traditional subjects such as: history, geography, physics, chemistry, biology, information science, etc.;
- achieving a complex unity between the informative and formative components of education;
- enlarging the action field of inter-disciplinarity and life-long education, etc.

The second approach will concentrate on:

- defining the social, psycho-pedagogical and methodological reasons for introducing new school-subjects required by Romania's new constitution (e.g. civics and democracy education, social welfare) and emerging market economy (e.g. economics, marketing, business and public administration);
- designing the objectives and contents of the new subject-matters, etc.

The design of the new curricula for Romanian pre-university education is the task of the newly-founded National Curriculum Commission made up of 3 subcommissions specialized in: Math and Science, Arts and Humanities, and Technical and Vocational Subjects.

References will be made to the main targets and work directions of these specialized groups of experts.

The theoretical framework of our paper will be completed with practical examplifications of curriculum strategies in the field of humanistic school-subjects, with special reference to foreign language teaching in secondary education.

CURRICULAR STRATEGIES FOR PROMOTING LIFELONG LEARNING

Eero Ropo, University of Tampere, Finland

The need for lifelong learning has become evident along the many changes taking place in the world around us. People have to acquire new sets of skills, abilities, and attitudes to live successfully in the changing societies. Life and work seems to require continuous learning. Learning has become a necessity and a lifelong process for everybody.

Past thinking in education has based on the idea that life can be divided into two parts: the parts of education and the work. New demands for lifelong learning challenges our thinking about the goals of education. Purpose of education according to the new ideology is to give people skills and abilities for lifelong learning, not so much knowledge that they can 'use'. All people are lifelong learners by birth, but their skills can and should be developed. We should help students to become expert learners. The purpose of the paper is to discuss theoretical issues in respect to designing new curriculums for promoting lifelong learning. The study is based on the theoretical analysis of current trends and ideas in the study of learning.

First, it is recognized that the concept of curriculum needs revision. It is necessary to include under it the contents and objectives of instruction, but also teaching methods, student experiences, etc. Most of the learning skills, for example, are learned as an outcome of the study process that different instructional methods initiate. In this framework didactic research on teaching methods, for instance, is considered as a part of curriculum theory. Thus, planning an educational environment to promote good studying and learning is an essential part of curricular planning.

As a second remark it is necessary to recognize that one of the most basic features of the current and future society seems to be its continuous change in all the sectors. Education for citizenship has to offer skills for coping with and directing these changes. To teach skills of both adapting to change and affecting with deliberate strategies to the directions of the changes, it is seen necessary to represent change processes in the curriculums. The paper will discuss current change processes affecting future goals and forms of education.

The paper will discuss five strategies that might be used when developing a curriculum promoting students' abilities for lifelong learning. First, the curriculum should be integrative. Knowledge, skills and abilities offered should integrate the person better into both the practical life and future education. This integration should be both horizontal and vertical. The education should increase an

individual's educational opportunities, not build artificial boundaries for development. Second, the curriculum should develop a learner's skills and abilities for new learning, thinking, and problem-solving in the real life situations. Lifelong learning can be seen as a skill to learn in formal, informal and non-formal education and the purpose of the curriculum is to promote the acquisition of the necessary skills. From a cognitive point of view an individual has to be able to set goals and objectives, direct, monitor and evaluate his/her own learning. Third, the curriculum should promote processes of personal growth and emancipation. This can be considered as one of the most important goals of education. To be a good lifelong learner an individual has to be an independent and emancipated person. This strategy sets new demands, for instance, for developing individual curriculums for persons in vocational and adult education.

In relation to the above strategies the paper will discuss some of the current ideas in the curriculum research, such as the complexity of learning environment, new medias for learning, and student participation into curriculum planning. The implications of the study for further research and educational practice will also be discussed.

EQUALITY THROUGH EDUCATION: AN EVALUATION OF INNOVATIVE PROJECTS CONCERNING EQUALITY IN SCHOOLS

Kirsten Reisby, The Royal Danish School for Educational Studies, Aarhus, Denmark

Abstract

The paper presents the results of an evaluation of innovative projects in the Danish Folkeskole (primary and secondary schools). The evaluation includes seven case-studies and is based on semi-structured interviews with teachers, pupils, and parents, and on material pertaining to the projects produced by the teachers.

The analysis of the data points to certain characteristics important to a pedagogy of equality, i.e. teaching which is supporting and challenging for both girls and boys,

- democratic decision-making processes,
- topics and themes which are open to accomodate for gender-specific experience,
- single-sex groupings as one way of organizing the class,
- training in group-processes and co-operation, and
- employing a variety of activities.

The theoretical base and the ideological assumptions of the evaluation as well as the innovative projects will be discussed in relation to further research and the development of a pedagogy promoting equity.

Summary

The evalutation wishes to contribute to the development of education for equality in theory and practice on the basis of a discussion of

- the intentions and the planning of the innovative projects
- the experiences of the teachers, pupils, and parents involved, and
- the results of previous research on gender and education.

The seven *innovative projects* in question were initiated by the teachers themselves. With regard to their focus they can be divided into two: three projects concentrated on teaching the "hard" subjects, Science and Technique to the younger children and Physics and Chemistry to the older age-groups. Four projects concerned the teaching of Danish and Social Sciences.

In evaluating these projects we used *the case-study method.* Each project was treated as a case. The *data* collected were: relevant papers, our own diaries of our visits to the schools, and semi-structured interviews with the respective headmasters, teachers, single-sex groups of pupils, and a selection of parents. With the teachers, the interviews often developed into dialogues.

143

The teachers were inspired by their own experience and by **research** pointing to the relational behaviour of girls and the achievement-oriented behaviour of boys. Girls are concerned with interpersonal relationsships. They are social, responsible, and good at co-operating. Boys are concerned with physical activities, with facts and with exploring the world. And they are mutually competitive.

The most significant characteristics of the pedagogical practice developed in the projects can be summed up as follows

- there is more emphasis on democratice decision-making processes because of girls' demands for participation
- topics and themes are open to accomodate for gender- specific experience
- there are periods with single-sex education in order to accomodate for gender-specific interests and to offer opportunities for exploring consciousness of own sex and gender-specific interaction
- group processes and co-operation are trained (mostly to meet the needs of the boys), and
- a variety of activities (manual, esthetic-creative, intellectual, etc.) and organizational forms are employed in order to break up the monopoly of classroom teaching, which has a tendency to favour the boys, and end the uniformity of traditional teaching with its emphasis on reading and verbal activities.

In addition to this, it is just as interesting to see that the educational intentions of the teachers have been changing over the last couple of years

- from viewing the competencies of the boys as the standard and therefore attempting to provide the girls with what they are lacking (e.g. interest, knowledge, and technical skills in Science)
- via an educational effort to support girls in the development of their sense of self-worth by giving them more space and attention in teaching situations
- to a promotion of equity aiming at giving each child optimal possibilities for developing a multiplicity of competences irrespective of sex, social background, handicap, ethnicity, and religion - within the context of democracy and solidarity. All the while considering sex and gender important variables.

What, then, has happened to effectuate this shift towards a pedagogy of equity/equal worth?

There is a tendency for the teachers to operate on two levels when considering their pupils. On a general level there are marked differences between girls and boys - differences explained in terms of biology as well as gender-specific socialization. The teachers wish to contribute to the dissolution of this pattern. On the level of individuals, however, the picture is more differentiated. Girls differ from each other, and so do boys, and few pupils are described as "real" girls and boys. The teachers wish to consider these differences in their pedagogical planning. Apparently the latter level is now prominent in the pedagogy for equity/equal worth. But why?

The following might be of importance:

Equality - equal worth is *personal, emotional, and - for some - an area of conflict.* In the pedagogy of equality the teacher's own experiences and emotions are activated, and they must be dealt with. This is so when female and male teachers co-operate, this is so in the confrontation with one's own sex or the opposite sex in

single-sex groups. This is so in relation to parent-teacher co-operation, where the pedagogy of equality can provoke emotional resistance.
In society at large there is an *ideological* concensus for equality, and at the same time there is also a widespread notion that equality has been realized. The teachers spearheading the innovation projects often meet with the attitude, that they are dealing with an obsolete problem. And they are suspected of personal motives (e.g. of having problems with their own sexual identity or sex role) for initiating the projects.
New *empirical* research points to the emergence of new "types" of girls and boys. *Theoretically*, this is explained by changes in the conditions for socialization and in the cultural signification of "gender" and the subsequent possibilities for new psychodynamic developments in both sexes.
A continued development of a pedagogy for equity/equal worth and the didactic consequences thereof must be considered in the light of the above.

References

An English version of the summary of the report : Gudrun Pedersen and Kirsten Reisby (eds) (1992), Equality and Multiplicity. On Equality in School. (The Royal Danish School for Educational Studies). Work in progress.

CREATING A CURRICULUM TO REDUCE STUDENTS' HIV–RISK BEHAVIORS

Sarah J. Stanley, IOX Assessment Associates, Los Angeles, USA

Introduction

The alarming spread of the human immunodeficiency virus (HIV) and the resultant numbers of deaths caused by the acquired immune deficiency syndrome (AIDS) has captured the attention of authorities worldwide. Because of the absence of an HIV vaccine, not to mention adequate treatment protocols for AIDS–afflicted patients, it is generally held that educational strategies must be employed to markedly reduce the behaviors that place people at risk of becoming infected. In spite of the widespread advocacy for potent HIV education, however, in at least some nations there has been less than a serious effort to provide high quality HIV education.

This is certainly the case in the United States, where the typical HIV education program provided in U.S. school districts consists of an hour or two of dominantly knowledge–focused instruction. Because there is ample evidence that such knowledge–oriented approaches have scant impact on students' actual behaviors, it is clear that more powerful HIV curricula must be employed.

A major effort to do precisely this has been funded by the Division of Adolescent and School Health of the U.S. Centers for Disease Control (CDC). One provision of a CDC contract with IOX Assessment Associates (IOX) called for the creation of a 15–hour HIV curriculum that can demonstrably reduce young peoples' engagement in behaviors placing them at risk of being infected with HIV.

This presentation will focus on the essential features of the new curriculum, created by IOX during 1990–1991. I will identify the distinctive elements that, based on related research regarding modification of other health–related behaviors, we regarded as being likely to prove efficacious. Educational researchers who have studied the impact of educational programs know all too well how difficult it is to bring about any significant changes in students as a consequence of an educational intervention. Even in routine attempts to enhance students' reading or mathematics skills, it is often difficult to get meaningful changes in students' pre–instruction to post–instruction performances. With respect to HIV education, of course, the obstacles to the modification of students' behaviors are even more profound. Because the bulk of HIV infection occurs as a consequence of people's engaging in sexual activity, the focus of most HIV education programs is to persuade students to abstain from sexual intercourse or,

if sexually active, to protect themselves by using condoms. The challenge is substantial.

Elements of Successful HIV Education Curricula

After an extensive review of previously reported investigations regarding the reduction of such risk behaviors as smoking, drug use, and unprotected sexual intercourse, the IOX staff arrived at the position that an effective HIV education program had to embody certain elements. These elements are listed below:

Attributes of Effective HIV Curricula

- The curriculum must require an instructional period of sufficient duration.
- The knowledge provided for students should be functional.
- Substantial instructional attention should be devoted to promoting students' vulnerability.
- Students' adoption of positive attitudes regarding their HIV–risk behaviors must be promoted.
- Students must acquire the kinds of interpersonal skills needed to reduce HIV–risk behaviors.

I will briefly consider each of these attributes and indicate how they were incorporated in an HIV curriculum, Skills for Life, created by IOX.

Adequate Duration

Is it any surprise that the typical 1–2 hour HIV education programs have little impact on students' behaviors? It is our view that a minimum of 10 hours of classroom instruction is necessary if the HIV educational program is to have any hope of modifying students' HIV–risk behaviors. The Skills for Life program was a 15–period program, offered during three consecutive weeks.

Ideally, an HIV education program will also incorporate the use of subsequent booster sessions of a few hours' duration to reinforce the instruction provided earlier.

Functional Knowledge

Student behavior is not likely to be influenced by knowledge about the physiological composition of the AIDS virus or the history of how HIV was identified as the cause of AIDS. Such general knowledge, while interesting, is not apt to alter students' HIV–risk behaviors. On the contrary, however, students' HIV–risk behaviors are far more likely to be altered by functional knowledge, that is, knowledge directly related to HIV–risk activity. To illustrate, there are certain facts associated with the acquisition, storage, and use of condoms that bear on whether condoms will provide individuals with the protection. An HIV curriculum should focus only on functional knowledge and should avoid ``nice to know'' content.

Vulnerability Perceptions

It is well known that most adolescents regard themselves as invulnerable to long term sickness, death, injury, and/or aging. Students who believe that they are invulnerable continue to engage in risky behaviors.

Accordingly, an effective HIV education program should incorporate instructional activities expressly directed at making students' perceptions of their HIV

vulnerability more accurate. For example, videotaped or in–person presentations could be made by adolescentswho had become HIV positive as a function of heterosexual activities. In the Skills for Life curriculum there were two half–hour videotapes employed that dealt specifically with sharpening students' vulnerability perceptions. In addition, an HIV–transmission ``game" was employed to illustrate how HIV can be transmitted among adolescents.

Appropriate Affect

Attitudes influence behaviors. Thus, an effective HIV curriculum should promote attitudes that are correlated with the avoidance of HIV–risk behaviors. For instance, the curriculum could foster sexual abstention by presenting attractive adolescent role models who recount, on videotape or in person, their own endorsement of sexual abstention until one is ready to engage in sexual intercourse. Similarly, positive attitudes regarding condom–use are particularly important.

The Skills for Life curriculum provides a number of videotaped individuals who endorse suitable attitudes and values regarding HIV–risk behaviors. In addition, there are specific activities designed to increase students' comfort levels regarding the use of condoms or engaging in less risky sexual activities.

Interpersonal Skills

Our review of research regarding relevant health–related educational interventions suggests that the acquisition of pertinent interpersonal skills (such as how to refuse a boy (or girl) friend's invitation to have sex without the running the risk of losing the boy (or girl) friend) is a particularly important ingredient in effective HIV education. If students possess interpersonal skills, such as how to negotiate the use of condoms, they tend to use these skills. In our view, an HIV education program should promote the following skills: (1) how to acknowledge a potentially risky situation, (2) how to escape from a risky situation, and (3) how to protect oneself in a risky situation one is unable or unwilling to leave.

Almost two–thirds of the Skills for Life 15–hour curriculum is devoted to the promotion of students' interpersonal skills. Students observe videotaped adolescents modeling the skills and engage in role–play applications of the skills.

Necessary, But How Sufficient?

In this summary I have set forth five key attributes that our review of relevant research suggests are needed if an HIV education program is going to influence students' HIV–risk behaviors. As suggested earlier, however, the modification of teenagers' sexual conduct is remarkably difficult. Empirical evaluations of programs such as the Skills for Life curriculum will indicate whether the five attributes identified here are sufficient. It may well be that we continue to underestimate what it will take to alter adolescents' HIV–risk behaviors.

SWEDISH NATIONAL ASSESSMENT PROGRAM FOR THE COMPULSORY SCHOOL

Oscar Oquist, National Agency for Education, Sweden

Objectives

- to assess the qualities of the Swedish compulsory school (age 7-16) in a broad spectrum related to the national curriculum (a total curriculum approach);
- to give the government and the public information on student achievments;
- to assess the impact of the school-culture on school results;
- to inspire local school development.

General premises

The assessment is based on the experimental round which took place in grade 2 and 5 during spring 1989. In this new round pupil data on all subjects are collected from grade 9 and to a limited extent (reading proficiency) from grade 2 and 5. Up to 15 lessons are to be devoted to the conduct of the tests. The instruments are developed by teams of researchers and practitioners at nine educational departments all over the country. Sample size is 10% in grade 9 and 3% in grade 2 and 5 each. The schools are selected by the pps-method (probability proportional to size). Total sample size is 16.000 pupils.

Perspectives

The results will be used to assess how well classes and schools attain the educational objectives and will also allow for analyses of variances between classes,schools and districts

Special importance will be attached to school unit level i.e. the level affording most opportunities for change.

A rich variety of instruments and methods should be used, including studies of non-cognitive variables,group-tasks and performance tests

The pupils are supposed to use their knowledge in functional settings

The results of the national assessment are to be available to all interests.

Method and techniques

- Area statistics;
- School unit descriptions (judgements by observers visting the schools);
- General questionnaires (to pupils,teachers and parents);
- Problem solving test (integrative);
- Subject-specific tests (n individual, pair and group settings).

Result and conclusions

The data have been collected in early spring 1992.Some preliminary results can hopefully be presented before summer. The first report will be presented in late autumn 1992.The results will be part of the report on "the state of the art" of the Swedish school system which NAE will present for the Government in autumn 1993. All schools are free to buy the instruments after the testing and use them as one mean for local evaluation.

Educational and scientific importance of the study

This is the first time such an encompassing assessment is made of the Swedish compulsory school("grundskolan").The program will show the entirety of the curriculum and the educational objectives.

The design of the program gives us an unique possibility to link achievment data to school-unit variables and more basic structural variables (multivariate methods of analysis).

The emphasis on qualitative data and process studies with teachers and pupils active in the evaluation process.

The decentralized organization of the assessment program. The close cooperation between different university departments will have important synergy-effects for test construction,data processing and analysis

References

Gustafsson, J.E. (1990). (Ed). Assessment of student outcomes as part of a national evaluation program. Methodological aspects and experiences from a first round. Göteborg University,Department of Education.

Truedson, L. What do they learn at school? Findings from a national evaluation of school. Stockholm, NBE 1990 (R 90:47).

TEACHER MATERIALS AS A PART OF COURSEWARE PACKAGES: CURRENT SITUATION AND DESIRABLE DIRECTIONS

Paul Keursten, & Nienke Nieveen, Department of Education, University of Twente, Enschede, The Netherlands

Introduction

Central to the implementation of computers in education is the classroom teacher (Office of Technology Assessment, 1988). Many teachers, if they are using computers at all, are still inexperienced users (Pelgrum & Plomp, 1991). In this initial implementation stage small scale successful experiences in their own classroom practice are an important factor in stimulating further use of the innovation (Fullan, Miles & Anderson, 1988). These experiences help teachers in acquiring clarity about the meaning and potential of the innovation, in gaining confidence in their own competence, and in developing their own view of the appropriateness of the innovation for the students and themselves.

Successful experiences in the initial implementation stage can only take place when teachers have high quality courseware at their disposal (Van den Akker, Keursten & Plomp, 1992). We define courseware as: an educational package, comprising at least computer software and written materials (for teachers and students). According to this broad definition, courseware is not restricted to drill-and-practice or tutorial applications. A courseware package can also consist of open ended software (like databases or word processors) and other (written) materials that induce classroom activities related to the software.

At the moment we are engaged in a five-year study in which we investigate which characteristics of courseware, especially teacher materials, contribute to successful implementation experiences in the initial implementation stage. In the first phase of this study design guidelines for courseware materials were formulated based on: an extensive analysis of empirical research literature; an analysis of the functions the materials should have for teachers; and three pilot projects. In the second part of the study we are testing the effectiveness of these guidelines in a large field experiment, in which an existing courseware package will be compared with an experimental version of the same package, developed in accordance with the design guidelines.

The literature review revealed that in research on the implementation of computer use in education, only limited attention is paid to teacher materials. Little is known about the current situation regarding teacher materials. For that reason we set up a special in depth study, focused on the following question:

- Do the currently available teacher materials within courseware provide teachers with adequate support to prevent or diminish implementation problems?

This paper will describe the design and outcomes of this study.

Method

From the available research literature it was apparent that teachers experience four central problems when implementing courseware. These are: the lesson preparation is complex and time consuming; a lack of background knowledge (leading to uncertainty); difficulties in changing their own role; an unclear view on learner effects (Van den Akker et al., 1992). These central problems were the starting point for a screening of currently available teacher materials that are part of courseware. To ensure the feasibility of the study, we limited the screening to packages intended for lower secondary education (age 12-15). We analyzed all courseware packages for lower secondary education that are currently on the Dutch market and that are intended for classroom use (167 packages).

For this screening of teacher materials we formulated a list of twelve questions, that were answered for each courseware package. Examples of these questions are: are the objectives of the package stated in the teacher materials; do the teacher materials contain specific advice for the preparation of the lessons; are there suggestions for monitoring student progress? This resulted in numerical information about the percentages of courseware packages for which a question could be anwered affirmatively. To gain more insight in how these characteristics are realized in teacher materials, we also quoted typical passages from packages in which the elaboration of one of these characteristics is extensive, about average, or minimal.

Results/conclusions

Data analysis was finished recently. Some general conclusions are:

- Currently, teacher materials provide very little support in diminishing implementation problems.
- Less than half of the packages contain procedural how-to-do-it advice for teachers. Particularly advice for lesson preparation and suggestions for monitoring learner effects are almost always lacking.
- In packages that do contain procedural advice, this advice is often limited to some general suggestions.

In the paper these (and other) results will be illustrated with typical passages from teacher materials.

Discussion

In discussing the results, we will relate the findings from this study to the more general literature about the role and functions of teacher materials in (curriculum) implementation. Based on this literature and on the preliminary results from our five-year study, we will discuss directions in which teacher materials could be improved.

References

Fullan, M.G. (1991). The new meaning of educational change. London: Cassell.

Fullan, M.G., Miles, M.B. & Anderson, S.A. (1988). Strategies for implementing microcomputers in schools: The Ontarion case. Toronto, Canada: Ontario Ministry of Education.

Office of Technology Assessment (1988). Power on! New tools for teaching and learning. Washington DC: Government Printing Office.

Pelgrum, W.J. & Plomp, Tj. (1991). The use of computers in education worldwide: Results from the IEA 'Computers in Education' survey in 19 educational systems. Oxford: Pergamon Press.

Van den Akker, J.J.H., Keursten, P. & Plomp, Tj. (1992). The integration of computer use in education. International Journal of Educational Research, 17 (1), 65-76.

OPEN LEARNING AS A CURRICULUM INNOVATION: A LONG ROAD FROM IDEA TO PRACTICE

E. van den Berg, Department of Education, University of Twente, Enschede, The Netherlands

Introduction

This paper reports about a study on the efforts made in curriculum development in three schools for adult education in the Netherlands. These schools have acquired special funds from the National Department of Education in order to establish an open learning program. Open learning is both a process which focuses on access to educational opportunities and a philosophy which makes learning more student centred. It is learning which allows the student to choose how to learn (modes of learning, e.g. lecture, seminar, independent learning, computer assisted instruction), when to learn (timing, frequency, duration), where to learn (classroom, library, at home) and what to learn (subject matter) (Cunningham, 1988, Paine, 1988). In an open learning system learning is less dependent on classroom instruction and more self-directed. This implies that the quality of (conventional and electronic) curriculum materials becomes far more important to steer the learning process.

Problem and methodology

The central problem in this study is how the process of curriculum development move forward. Special attention was paid to potential pitfalls of schoolbased curriculum development within the context of a large scale innovation project.

The typology of Goodlad et al; (1979) for different manifestations of the curriculum - ideal, formal, perceived, operational and experiential - guided the data collection and analysis.

An in-depth case study has been carried out with two stages of data collection in the school year 1989-1990 and 1990-1991. In these periods the process of curriculum development and implementation were documented, using interview and observation procedures and analysing documents produced by the schools. After the first raw data analysis the results were compiled in a summary report (Miles & Huberman, 1984).

Results and conclusions

The designing of the open learning curriculum consisted of two stages. Firstly, a curricular framework was developed and, afterwards, based on this framework, the instructional materials were developed.

In designing the curricular framework there were three main problems.

In the first place there was not a shared vision between school management and teachers about open learning. The concept clarification or ideal curriculum was very weak. Van der Vegt and Knip (1988) pointed out that it is essential to formulate a concept of the desired state of affairs that serves to focus attention and energy. Also Walker (1990) stipulate the importance of a 'platform of ideas'. In the experimental schools this was not the case.
Secondly, there was hardly any time to communicate in a formal and informal way about the innovation, because of much part-time faculty.
Thirdly, there was no adequate training to support the design process and no systematic design model was used.
As a result a curriculum has been developed in which the traditional program was divided into periods of eight weeks, but the program has no flexibility in the way that students can choose how they will plan and conduct their learning activities. So the linkage between the ideal curriculum (an open learning program) and the formal curriculum (a traditional curriculum outline) was inadequate.
The development of the instructional materials was not steered by an acceptable curriculum outline. Moreover only one or two teachers were in charge with the development of a learning package. These teachers did hardly get any instruction or guide-lines to support them. As a consequence the authors, also because of time constraints, relied heavily upon the traditional textbooks. Only printed materials were produced and these materials were implemented without formative evaluation. Many teachers reported severe problems when they actually used the instructional materials.
Interviews with teachers and classroom observations showed that the poor material were the main cause of the failure in implementing an open learning program. Another important reason was that many teachers and students were sceptical about the idea of open learning.
We can conclude that both management and teachers, who were authors, underestimated the time, finances and know how needed for designing and developing an innovative curriculum resources.

THE LINKING OF CURRICULUM MATERIALS DEVELOPMENT WITH IN-SERVICE TRAINING FROM AN IMPLEMENTATION PERSPECTIVE

M.G. Roes, Department of Education, University of Twente, Enschede, The Netherlands

Introduction

In the Netherlands secondary schools are facing important curriculum changes. Currently, students at the age of 12 have to make a choice between different types of secondary education that differ with respect to level, focus (general or vocational) and subjects taught. The current curricula, which have been in use since 1968, are to be replaced by a curriculum consisting of 15 subjects for all students in the age of 12-15. Minimum goals are set for these 15 subjects. The structure of the secondary school system stays intact.

The National Institute for Curriculum Development (SLO) has been charged with developing examples of the new curriculum at national, school and classroom level. To stimulate the use of those examples, SLO is planning to cooperate with institutions for in-service teacher training. A basic assumption for this cooperation is that curriculum materials, especially at classroom level, can have an important role in in-service training.

The research project "Curriculum development and in-service training" (LENA) is aimed at the following question:

- What characteristics of curriculum materials at classroom level contribute to more and effective use of those materials in in-service training?

Theoretical background

In recent years a lot of knowlegde about factors that affect implementation has become available. Fullan (1985) has summarized this knowledge and mentions as crucial factors:

- the development of clear and validated materials;
- focused, ongoing in-service training;
- interaction and collegiality at the school level;
- active leadership and administrative support.

The combination of materials development and in-service training is potentially very effective in efforts to improve educational practice. An obvious link between curriculum development and in-service training is the use of exemplary curriculum materials as a source and tool for in-service activities. The various functions of those materials in in-service training can be derived from knowlegde about

effective in-service training practices, for instance from the model of Joyce and Showers (1988). This model consists of the components theory, demonstration, practice, feedback and coaching. Van den Akker (1992) gives the following elaboration:

- theory
  The materials can present background information about the rationale of the curriculum innovation.
- demonstration
  The protocols in the materials can serve as tools for demonstration of both the ideal performance and the potential problems in the teaching practice.
- practice
  The materials can help teachers in practising new skills by providing specific guidance for classroom activities in exemplary lessons.
- feedback
  Teachers can receive feedback on their teaching activities by peer-teachers, using the same materials and exchanging their experiences, or by external observers who have expertise on the innovation in question.
- coaching
  The materials can also be used in follow-up activities when teachers try to implement the new skills in their own instructional setting.

In this research project the function and characteristics of curriculum materials in in-service training will be further explored. Besides a scenario will be developed for the transfer of exemplary curriculum materials to the institutions for in-service training.

Method

First, an inventory will be made of: a) the exemplary curriculum materials SLO has developed/is developing with respect to the innovation and b) the in-service activities that are planned with regard to the curriculum change. This inventory will be made by means of document analyses and interviews. The results of the inventory will be used for the selection of at least two pilot projects in which a scenario will be developed for the transfer of curriculum materials to the institutes for inservice-training.

Results

At ECER, the results of the inventory will be presented together with the research design for the next stage of the project.

References

Akker, J.J.H. van den (1992). Die Grundschule unterwegs: Ergebnisse aus der Innovationsforschung anderer OECD-Länder. In W. Hameyer, R. Lauterbach & J. Wiechmann (Hrsg.), Innovationsprozesse in der Grundschule. Fallstudien, Analysen und Vorschläge zum Sachunterricht (pp. 306-316). Bad Heilbrunn: Klinkhardt.

Fullan, M. (1985). Curriculum implementation. In T. Husen & N. Postlewaite (Eds.), The International Encyclopedia of Education (pp. 1208-1215). Oxford: Pergamon Press.

Joyce, B., & Showers, B. (1988). Student achievement through staff development. New York: Longman.

THE USE OF CLASSROOM CONSULTATION IN IMPROVING THE IMPLEMENTATION OF A CURRICULUM

J. Snippe, Institute for Educational Research (RION), Groningen, The Netherlands

Introduction

It is a well known problem that the implementation of innovations, especially curricula, is often very poorly. This is partly due to the fact that teachers do not use the curriculum in the way it was meant to be. In literature several techniques have been described to improve the implementation of a curriculum by changing the teachers' behaviour through some sort of in-service training. Organizing in-service meetings or applying micro teaching are widely used interventions. One of the more promising interventions is the use of classroom consultation (coaching). This is a form of individual support for teachers in the classroom. Teachers are observed while they teach by a consultant who later discusses the lesson with the teacher. In this feedback-conversation discreacies between the observed and the desired behaviour are discussed and recommendations are made for improvement. Several experiments have shown that classroom consultation can have an effect on teachers' behaviour and on pupils' achievement (Showers, 1982; Stallings, 1985; Snippe, 1991).

In Holland the support for teachers is one of the main tasks of the Teacher Advisory Centres, local and regional school support agencies. The main question is what interventions these centres use to support the implementation of innovations and to what extent they make use of classroom consultation. Another question of importance is the way teachers and principals look upon the use of this intervention.

Method

In this paper the results of two research projects wille be reported, that have taken place at two Teacher Advisory Centres. The projects both have a correlational design. One of the questions in these projects concerns the use of interventions in these centers to affect implementation of curricula, particularly the use of classroom consultation.

In the first project a questionnaire was filled in by 39 school principals, 90 teachers from elementary schools and 14 teacher advisors. The questionnaire contained questions concerning the amount and sort of support teachers and principals achieved and how they experienced this form of support.

The second project was conducted to evaluate the results of an implementation project, in which one teacher advisor supported the implementation of a math curriculum in 21 elementary schools. The interventions that were used were self-

evaluation forms, classroom consultation and in-service meetings. A questionnaire was filled in by the teachers who were involved in the project.

Results

The results from the first project showed that teachers receive very little individual support in the classroom while implementing innovations, such as new curricula. Those teachers who have received classroom consultation are not very positive on the results: they don't believe it has contributed to a great extent to solve their problems. Change agents also say they give little individual support to teachers, although they also indicate that teachers have a need for classroom consultation. This has not been confirmed by the teachers. There is also very little support for principals who want to learn to coach their teachers. The principals indicate they need this sort of support. When they receive support in order to learn to coach they indicate it helps them solving their problems as an internal facilitator of change. The results from the second project will be available in march this year.

Discussion

The results as far as they are available at this moment, show that although the value of classroom consultation has been empirically established to a large extent, the implementation of this intervention in educational practice meets a lot of problems. One of these problems is the acceptance by teachers. When classroom consultation can not offer teachers an acceptable and concrete solution to their implementation problems then teachers will not value this intervention very high. On the other hand teachers may react this way out of a sense of threat. Classroom consultation can be a means of diminishing the feeling of isolation many teachers have in their team.

References

Showers, B. (1982). Transfer of training: the contribution of coaching. Eugene: Beverly Oregon Univ., Eugene Center for Educational Policy and Management.

Snippe, J. (1991). In-service training voor leerkrachten. Proefschrift. Groningen: RION.

Stallings, J.(1985). A study of implementation of Madeleine Hunter's model and its effects on students. Journal of Educational Research, 78, 6, 325-337.

ACTIVITY-BASED LEARNING IN ELEMENTARY SCHOOLS: IMPLEMENTING A CORE CURRICULUM PRINCIPLE IN ELEMENTARY SCIENCE EDUCATION

U. Hameyer, Institute for Science Education (IPN), Kiel University and School of Education, Germany; J. van den Akker, Department of Education, University of Twente, Enschede, The Netherlands

A comparative cross-country analysis of curriculum renewal will be discussed. The study is called IMPACT which is the acronym for »Implementing Activity-Based Learning in Elementary Science Teaching«. IMPACT was carried out in the Netherlands, Sweden, United States and Germany. In this project we studied lasting processes of curriculum renewal in the realm of activity-based learning in elementary science teaching of 9 and 10 year old children. The methods we used are based on qualitative data analysis and comparison of 15 in-depth case studies. We mapped out what productive schools have in common, for instance sufficient capacity to establish good climates and conditions for putting activity-based curriculum ideas into practice. The contribution will focus on the methodology of IMPACT, the conceptual model called PROMISE (Process Model of Institutionalization and Self-Renewal), selected preliminary findings, and a structural look at basic conditions necessary to facilitate lasting curriculum renewal.

Lasting school development currently gains growing interest among curriculum researchers. Classical studies often investigated short-term changes covering a few months. Explorative studies, longitudinal designs and sound theories are needed to analyse the school curriculum as being practiced over time spans of more than a year or two (see Miles/Ekholm/Vandenberghe 1987).

Our contribution will focus on results from an international study called IMPACT which is the acronym for »Implementing Activity-Based Leaning in Elementary Science Teaching«I. The findings and case studies will be available in 1992 (Hameyer/Akker/Anderson/Ekholm 1992/93 and Ekholm/Hameyer 1992/93).

In IMPACT, we studied *lasting processes of* curriculum *renewal* with emphasis on conditions under which activity-based learning in elementary science teaching of 9 and 10 year old children became part of routine working patterns in effective schools. The methods we used to study processes of institutionalization which encompass 3 to 7 years are based on *qualitative data analysis* and *systematized cross-case comparison.*

156

It is particularly revealing to explore productive schools from inside and to learn from their *organizational biography*. In IMPACT we comparatively map out what productive schools have in common, for instance sustained capacity to establish good working climate, cooperative efforts, and other institutional requirements for putting activity-based curriculum ideas as a common principle of basic education into lasting practice.

We describe practice profiles of 15 schools from four countries (US, the Netherlands, Germany and Sweden). Exemplary lessons are studied which emphasize activity-based learning, and we add empirical information about the degree to which students organize their learning activities on an increasin~gy autonomous level. Further we investigated how principals encouraged the process of curriculum development and to what extent the environment such as parents, authorities, support personnel, and educational policy-makers stimulate or impede the process.

The IMPACT cases have been selected and analyzed according to a framework which we were able to elaborate. The result is a conceptual model called *PROMISE - Process Model of Institutionalization and Self-Renewal*. This model will be explained further in the contribution and illustrated by results and case study vignettes. We will focus on the curriculum aspect of IMPACT which itself also covers organizational factors of change processes on the school level.

The curriculum aspect of IMPACT shows up that activity-based learning gained strong emphasis in all our schools with the aim to enrich experiential knowledge and meaningful learning. The schools we studied partly reshaped their curriculum practice profile by new patterns of teaching professional staff development, school-based cooperation, and creating stimulative learning environments which allow for students' active experimentation.

Note 1

The research was conducted in Germany, Sweden, The Netherlands, and the United States of America (see *Hameyer/Wiechmann* 1991) The framework emerged from interdisciplinary cooperation of researchers from the following institutions: Kiel Institute for Science Education, Germany, University of Twente, the Netherlands, University of Colorado, USA, and University of Göteborg/Karlstad, Sweden.

EXEMPLARY TEACHING PRACTICES AND SCIENCE SUBJECTS CURRICULUM REFORM

W.A.J.M. Kuiper, Department of Education, University of Twente, Enschede, The Netherlands

Introduction

In the second quarter of 1989 a precoded written questionnaire was submitted to a stratified random sample of 1184 physiscs, chemistry, biology and general science teachers in lower secondary education (age 12 through 15/16) in The Netherlands. Purpose of this survey was to supply representative, descriptive research data on the actual teaching practice. Data were collected on a.o.:

- use of textbooks and other curriculum-materials;
- the science subjects instruction topics;
- the science subject instructional process (a.o. relation between the content of science subject instruction and real-life phenomena/applications students are acquainted with);
- mode of presentation and student engagement;
- laboratory and field work.

One of the results was that as grades rise biology, physics and chemistry teachers place less emphasis on real-life phenomena and applications students are acquainted with and pay more attention to formal subject-specific concepts and principles (Kuiper & Alting, 1990).

The survey was the *first part* of the so called Science Subjects Assessment Study (SSAS) and resulted in empirical base-line information for the implementation of 'basic education'. Basic education refers to the reform in view of lower secondary education. The most important characteristics of this reform are:

- the implementation of a core curriculum of 15 subjects - a.o. physics/chemistry and biology - for both lower general and vocational education (age 12 through 15), starting from August 1993;
- standardised attainment targets for all 15 subjects;
- no changes in the currant categorial structure of lower secondary education.

The most important aim of basic education is a curriculum reform. As far as the science subjects (physics, chemistry, biology) are concerned, the necessity is urged to relate subject-matter to real-life phenomena and applications students are acquainted with (so called 'contexts'). Besides there ought to be more opportunity for less receptive, more activity-based student engagements (especially conducting experiments). According to many such a reform of the science subjects curricula is necessary in order to promote the formation of

meaningful concepts, to facilitate the application of concepts to realistic situations and to make instruction more attractive and motivating for students (cfr. De Lange, 1987). The curriculum reform in view has its roots in the constructivistic theory of learning and instruction (Jonassen, 1991, 1992; Tobin, Kahle & Fraser, 1990; Tobin & Gallagher, 1987). General conclusion from the survey is that there is a considerable gap between the ideals (the curriculum reform in view) and the actual teaching practice (Kuiper & Alting, 1990).
This paper describes the objectives, design and results of the *second part* of SSAS. Purpose of this part of the study is an explorative description and analysis of the teaching practices of some exemplary science subjects teachers. We know curriculum reform is a complex, refractory and multidimensional process, demanding the possible use of new or revised materials, the possible use of new teaching approaches and the possible alteration of beliefs (Fullan, 1991). The assumption on which the second part of SSAS is based, is that the study of exemplary teaching performances of experts - by focusing on successes, exciting experiments and positive facts (Pennick & Yager, 1983) - could provide support to the improvement of the science subjects teaching practices in the direction of the curriculum reform in view (cfr. Berliner, 1986; Pennick & Yager, 1986; Shulman, 1986; Tobin & Fraser, 1987). Recently several small-scale research efforts on exemplary science teaching were inititated in the USA, Australia and Israel (a.o. Ben-Zvi, Hofstein & Carmeli, 1990; Fraser, Treagust & Tobin, 1990; Gallagher, Bien & Karunaratne, 1990; Tobin & Fraser, 1987; Tobin & Fraser, 1990; Yager & Bonnstetter, 1990).

Research question
The central research question of the second part of SSAS is: Wich recommendations for the development and implementation of context-bound and activity-based science subjects education (in the period of basic education) can be distilled from the teaching practices of some exemplary teachers? The study of the exemplary teaching practices focuses on (cfr. Fullan, 1991):

- curriculum-material (kind, use, the teachers' perception of its quality and practicality);
- teaching approaches (use of contexts, opportunity for activity-based student engagements);
- the teachers' opinions on the curriculum reform in view (need, clarity, complexity).

Primary focus of interest is what actually happens in the science subject classroom (cfr. Tobin & Gallagher, 1987). An important objective of the study is the formulation of specifications for designing context-bound and activity-based curriculum-materials.

Design, instruments and procedures
Between October 1989 and June 1990 8 case studies (1 pilot not included) were conducted, focusing on the exemplary teaching practices of 4 biology and 4 physics teachers in lower secondary education. The teachers were selected from the 856 teachers who completed the survey-questionnaire. They were selected because, as appeared from the survey, they obviously tried to relate the content of science subject instruction to the world outside the classroom and spent a substantial amount of time to laboratory work.

Data were collected by means of:

- *direct observation* of lessons, per case in one group per relevant grade; altogether 76 lessons (34 physics, 42 biology) were observed; the course of the observed lessons was recorded on paper;
- *textbook analysis* (checklist);
- an extensive *interview* with each teacher at the end of the observations; topics were the course of the observed lessons (use of contexts, opportunity for activity-based student engagements), use of and opinion on the textbook involved (quality and practicality), and opinion on the curriculum reform in view (need, clarity, complexity);
- an *analysis* of some *achievement tests* (developped by the teachers themselves);
- submission of a precoded *questionnaire to students* (per case all groups in all relevant grades); the questionnaire contained precoded items on students' perception of the operational curriculum (teacher approaches, student engagements), students attitudes to the science subject involved and students' opinions on the textbook involved; altogether 865 student questionnaires were completed (294 physics, 571 biology; pilot not included).

All case studies were conducted by one and the same researcher. Data were reported per case. Consequently each case study report was sent to the teacher involved for validation.

Results, conclusion and discussion

All 8 case study reports have been approved by the teachers involved. The data are analysed at the moment. The results of the observations, interviews and textbook analysis will be described and discussed in the paper.

References

Ben-Zvi, R., Hofstein, A. & Carmeli, M. (1990). The influence of exemplary teaching on students' enrollments in and perceptions of chemistry studies. Paper presented at the 1990 NARST Annual Meeting, Atlanta, Georgia.

Berliner, D.C. (1986). In pursuit of the expert pedagogue. Educational Researcher, 15 (7), 5-14.

Fraser, B.J., Treagust, D.F. & Tobin, K. (1990). Australian studies of exemplary science teaching. Paper presented at the 1990 NARST Annual Meeting, Atlanta, Georgia.

Fullan, M.G. (1991). The new meaning of educational change. London: Cassell Educational Limited.

Gallagher, J.J., Bien, L. & Karunaratne, S. (1990). Science teaching in two Michigan elementary schools. Paper presented at the 1990 NARST Annual Meeting, Atlanta, Georgia.

Jonassen, D.H. (1991). Context is everything. Educational Technology, 31 (6), 33-34.

Jonassen, D.H. (1992). Objectivism versus constructivism: do we need a new philosophical paradigm? Educational Technology: Research & Development, 39 (3), 2-20.

Kuiper, W.A.J.M. & Alting, A. (1990). Biology, physics, chemistry and general science: the actual teaching practice. Enschede: Universiteit Twente, OCTO (in Dutch).

Lange, J. de (1987). Mathematics, insight and meaning. Utrecht: Rijksuniversiteit, OW&OC.

Penick, J.E. & Yager, R.E. (1983). The search for excellence in science education. *Phi* Delta Kappan, 64, 621-623.

Penick. J.E. & Yager, R.E. (1986). Trends in science education: some observations of exemplary programmes in the United States. European Journal of Science Education, 8, 1-8.

Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. Educational Researcher, 15 (2), 4-14.

Tobin, K., Butler Kahle, J. & Fraser, B.J. (1990). Learning science with understanding: In search of the holy grail? In K. Tobin, J. Butler Kahle & B.J. Fraser (Eds.), Windows into the science classrooms. Problems associated with higher-level cognitive learning (pp. 1-13). London: The Falmer Press.

Tobin, K. & Fraser, B.J. (1987). Exemplary practice in science and mathematics education. Perth: University of Technology.

Tobin, K. & Fraser, B.J. (1990). What does it mean to be an exemplary science teacher?, Journal of Research in Science Teaching, 27 (1), 3-25.

Tobin, K. & Gallagher, J.J. (1987). What happens in high school science classrooms? Journal of Curriculum Studies, 19 (6), 549-560.

Yager, R.E. & Bonnstetter, R.J. (1990). The practices of teachers who develop exemplary science programs. Paper presented at the 1990 NARST Annual Meeting, Atlanta, Georgia.

THE EXPERIENTAL CURRICULUM: CONCEPTUALIZATION AND MEASUREMENT OF DIFFERENT LEVELS OF PUPIL EXPERIENCES

R.F.A. Wierstra, Department of Educational research, University of Utrecht, Utrecht, The Netherlands

Introduction

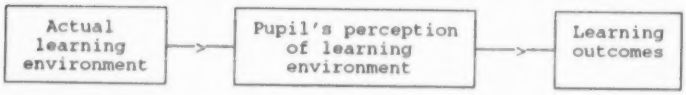
Goodlad et al. (1979) make a distinction between several kinds ('domains') of curricula. Among these are the formal curriculum, the operational curriculum and the experiental curriculum. As Van den Akker (1988) noticed these concepts are not unambiguously defined. In this paper we focus on the experiental curriculum. In terms of Goodlad et al. this construct refers to the curriculum as actually experienced by students.

The concept 'experiental curriculum ' is in our opinion rather ambiguous, because at least three conceptually different levels of pupil experiences may be discerned: pupil perception of the operational curriculum (we will indicate this level as learning environment perception); pupil appreciation of the curriculum, and, on a higher level, affective outcomes (generalized subject matter attitudes) which result from the curriculum. For all three constructs, pupil questionnaires can be used, but the constructs behind the questionnaires, and consequently the questionnaire designs and the ways of validating the questionnaires, are different.

In this paper the conceptualization and measurement of the three constructs will be illustrated within the framework of evaluative research. All three experiental constructs are important for curriculum evaluation, especially for causal–analytical evaluation as advocated by Van den Berg, Terwel & Wierstra (1992), but are not used very often by researchers. Especially measurement of learning environment perception has not been taken seriously by researchers. This in spite of Walberg's claim that a pupil's perception of the learning environment explains the only sizeable variation in (immediate) learning outcomes beyond that acounted for by aptitude and 'content opportunity'(Walberg, 1976). Walberg, therefore, adheres to a mediation hypothesis in that he considers perception as a molar judgement that 'may mediate the multiplicity of molecular events of instruction and other classroom activities and properties...Perception can usefully and simply index the complex match of internal and external elements, structures and sequences that optimizes learning'(Walberg, 1976, p. 160; emphasis is ours). This mediation hypothesis is represented diagrammatically in figure 1.

Fig. 1

A Simplified Version of Walberg's Mediation Hypothesis



Experiental constructs in curriculum evaluation

The evaluative research that will be discussed focuses on a comparison of an innovative Dutch physics curriculum (the PLON curriculum), with more conventional physics curricula. The relations between the experiental levels were investigated in the context of the following evaluation questions:

- To what extent do pupils who follow the PLON curriculum experience the learning environment as more reality-centred and more pupil-centred than pupils from conventional classes?
- To what extent do pupils from PLON classes appreciate physics lessons, to what extent do they attain a more positive attitude to physics, and how are these factors related to the classroom reality–centredness and pupil-centredness, and to pupil's perception of these aspects ?

Reality-centredness and pupil-centredness are characteristics of the learning environment as intended by the PLON curriculum developers. Reality-centredness refers to the extent to which the subject of physics is presented explicitly in relation to everyday life and to the out-of-school experiences of pupils. Pupil–centredness means an emphasis on activity–learning and on taching–learning procedures which allow pupils to influence the teaching–learning process.

Method

The investigation was carried out among fourth–grade HAVO pupils, who followed instruction in mechanics. A pretest-posttest control group design was used for 669 students from 25 classes.

Pupil perception of the learning environment (reality-centredness and pupil-centredness) was measured after the unit by means of a questionnaire consisting of 34 statements of the following kind:

'All of the pupils were engaged in the same activities at the same moment'; 'In the physics lessons situations out of school were dealt with, which showed what physics had to do with daily life. Response alternatives were: (almost) never,

seldom, sometimes, often, very often. A similar questionnaire was given to the teachers.
In addition to the learning–environment measurements we measured pupil appreciation of the mechanics lessons and the experienced instructiveness (nature and intensity of subjective learning experiences). At the beginning of the unit these aspects were also measured with regard to physics lessons in general. In addition, pupils' attitudes to reality-centred and pupil-centred instruction at the beginning of the unit were measured.
All scales proved to be reliable. For example, high Cronbach alpha values were obtained for the learning –environment measurements (.80 or higher for each curriculum group) as well as high IRT reliability (Mokkenscale reliability, Gustaffson ISS index).
Research question 1 was answered by means of ANCOVA, nesting pupils within class within curriculum; The analyses with regard to research question 2 comprised among others multilevel analysis (using as predictors initial knowledge, initial pupil attitudes, curriculum, and learning environment perception, the latter both on pupil level and on aggregated class level).

Results
The learning–environment scales were also very reliable on class level, as indicated by intraclass-correlation coefficients (Winer, 1971). There were also strong correlations for these scales between aggregated class scores and teacher scores. These data suggest that the individual learning–environment score partly indicates the real learning environment. This suggestion is confirmed by the fact that pupil learning–environment scores turned out not to be correlated with pupil characteristics such as gender and initial attitudes (initial attitude to physics, to reality-centred instruction and to pupil-centred instruction). There were, however, high correlations in each curriculum group between learning environment scores and affective outcomes (appreciation and perceived instructiveness). These correlations were higher for individual learning–environment scores than for aggregated scores. This is in agreement with Walberg's mediation hypothesis (Walberg, 1976) and with our former results (Wierstra et al., 1986).
With regard to evaluation–question 1 it was found that the PLON group experienced a more pupil-centred and a more reality-centred learning environment than the conventional group. With regard toresearch question 2 the answer is more complex. The PLON group reported clearly more reality-oriented learning experiences than the conventional group.

Some concluding remarks about the experiental constructs
In spite of the subjective, individual components of the learning –environment measurements, we may conclude that the aggregated learning–environment score for a class gives a picture of the actual learning environment as it might be perceived by a trained non-participant observer. In this context Goodlad et al. (1979) declare: 'what goes on hour after hour, day after day in school and classroom. There is no way of knowing, for sure, what it is. The operational, too is a perceived curriculum'. The data indicate that pupils can give a description of the operational curriculum too. At present we are investigating how strong the correlation is between the questionnaire data and observer data.

A distinction should be made between student perception of the curriculum (the individual learning environment score) and two kinds of affective student response to the curriculum. The present study presents empirical confirmation of this distinction. It also demonstrates that the distinction does not exclude relations between the three levels. Perhaps teachers and instructional designers should be even more aware of these relations in future.

References

Akker, J.J.H. van den (1988). Ontwerp en implementatie van natuur–onderwijs. Dissertation. Lisse: Swets & Zeitlinger.

Berg, G. van den, Terwel, J. & Wierstra, R.F.A. (1992). Evaluatie van leergangen. Lisse: Swets & Zeitlinger (in press).

Goodlad, J.J., Klein, M.F. & Tye, K.A. (1979). The domains of curriculum and their study. In J.L. Goodlad e.a., Curriculum inquiry: The study of curriculum practice. New York: McGraw–Hill, 43–76.

Walberg, H.J. (1976). The psychology of learning environments: behavioral, structural or perceptual? In: L.S. Shulman (Ed.), Review of Research in Education, 4, 142–178.

Wierstra, R.F.A., Jörg, A.G.D. & Wubbels, Th. (1987). Contextual and individually perceived learning environment in curriculum evaluation. In: B.J. Fraser, The study of learning environments, 2. Perth: Curtin University of Technology.

Winer, B.J. (1971). Statistical principles in experimental design.New York: McGraw Hill.

FORMATIVE RESEARCH IN CURRICULUM STUDIES

Decker F. Walker, School of Education, Stanford University, Stanford, USA

In advanced industrial societies curriculum decision systems are breaking down under the strain of rapid, extreme technological, economic, and social changes. The main symptom of breakdown is the prevalence of curriculum reform movements centered outside the established decision systems that seek to circumvent and overthrow them in favor of direct political action. The ideas that inspire curriculum reform movements are usually sweeping, dramatic, and ideologically charged — radical new forms of education or radical returns to earlier forms. For advice on how to realize their vision on a wide scale, reform leaders look to prominent scholars who have written about the vision and to practicing educators who have succeeded in creating programs similar to those called for in the vision. They seldom look to the research and scholarly community as a whole or to the body of knowledge that community may have accumulated over decades of studying curriculum phenomena and problems. Thus curriculum studies have little influence on curriculum reform efforts, a situation that I believe is bad not only for the reform movements and the field of curriculum studies, but also for education systems and the societies they serve.

The main point of this address is to propose that forms of inquiry particularly suited to inform and guide curriculum change efforts belong at the center of the field of curriculum studies. Curriculum theory already provides ideas to guide reform, and curriculum criticism explores the relationship of these ideas to other ideas and values and to plans, policies, and practices. But theory and criticism cannot do their best work unless they are rooted in experience. It is not enough for curriculum studies as a field of professional inquiry to conjure up wonderful visions of higher curricular possibilities and to criticize such visions and the programs drawn from them. We must also show that the promises implied by these visions can be kept, that programs matching the vision can be realized in some real settings. And we must discover what are the costs and risks, as well as the advantages, of programs that realize the vision. Without such information, which can only be drawn from experience, those who shoulder responsibility for curriculum reform have only ideals to guide them, and the pages of history are littered with disasters wrought by high-minded idealists hell-bent on pursuing their visions heedless of costs and risks.

The main idea behind this proposal was first expounded by John Dewey in his 1929 book, *Sources of a Science of Education.* Dewey called it the *test of practice* and argued that no educational idea or principle can be said to be educationally

sound (scientific in Dewey's day) until it has passed the test of practice. In its simplest terms, passing the test of practice means that first, these ideas and principles have been embodied in programs and practices that are well enough understood to be described, and recreated in recognizable form by others in other settings; second, that these programs have been made to work in an appropriate range of real educational situations; and, third, that those whose judgments count in real educational decisions find the results of using these programs to be more educative, on balance, the results obtained by using the other programs and practices available.

The course of action I propose for researchers and scholars in the curriculum field is to make the test of practice the foundation of our empirical work. To do this means, I think, to make *formative research* one of the staple forms of inquiry in the field. The notion of formative evaluation is familiar - to assess the worth or value of a product or program so that it can be improved. Formative research pursues the same goal but begins at an earlier stage. Formative research is research carried out to understand an educational problem more clearly and deeply, to explore and weigh alternative avenues toward its solution, and to assess the ideas, strategies, and tactics that may one day yield a product or program to be formatively evaluated.

A variety of methods of inquiry are appropriate for formative research, including some that are not well-known or widely used in educational research today. The methodological criteria and standards commonly applied to empirical research in education will not be entirely appropriate for formative research. New institutional arrangements will probably be needed for supporting formative research and for communicating its results. These and other aspects of formative research will be discussed in the address.

DATABASE TECHNOLOGY FOR CURRICULUM COMPARISON AS APPLIED TO INFORMATICS EDUCATION

Fred Mulder, & Anneke E.N. Hacquebard, Department of Informatics and Information Technology, Dutch Open University, Heerlen; Deny Smeets, Hogeschool Gelderland, & Tinie Veenstra-Strijland, Nijenrode, The Netherlands School of Business, Breukelen, The Netherlands

Abstract

In order to make an adequate comparison of curricula it is necessary to map their subject descriptions onto a standard terminology. Aiming at such a standard for informatics education, a classification scheme has been developed on the basis of existing systems for classifying literature and some recently proposed model curriculum descriptions. The methodology for deriving this so-called Unified Classification Scheme for Informatics Education (UCSIE) draws heavily and successfully upon database technology. And so does, naturally, the storage and processing of the large number of the individual curricula to be compared. 'Databases as essentials in the curriculum field', that is the central theme of this paper.

Introduction

Informatics is of growing importance in a variety of educational programs. Also the contents of 'full' informatics curricula has changed rapidly in the last decade. When having such a dynamic educational field it is worthwile to get an overview of the differences and similarities between curricula. It is the ultimate goal of the research project reported in this paper to obtain such an overview for the Dutch informatics education. Extensive information on a full range of informatics educational programs has been acquired and stored into a database in order to allow for a comparison with respect to the weight and the level of the different informatics components, the study load for each subject, the scheduling of the subjects, the educational context and methods, etcetera. Such a comparison can only be made in a consistent and transparant way if the subject descriptions of the specific programs are all mapped onto one and the same terminology. Part of the research project was to develop a standard terminology, contained in the so-called Unified Classification Scheme for Informatics Education (UCSIE); see Mulder et al., 1992. Another part, not yet finished, is to use UCSIE in the actual comparison of curricula.

Method

The first task was to develop a classification scheme for informatics education. Various systems for literature classification have been studied, such as the Computing Reviews Classification System (Coulter, 1991) and the system used by the journal Computer Abstracts. In this study we have also included the description schemes of a number of recently proposed model curricula, for example the ACM/IEEE-CS curriculum (ACM/IEEE-CS, 1990; Tucker et al., 1991), which is the first attempt to 'join' the two traditionally separated fields of 'computer science' and 'computer science and engineering'; another example is the DPMA-curriculum for 'information systems' (DPMA CTF-90, 1990), yet another area apart but also related to the two fields mentioned before. From all these sources a new three-level classification scheme has resulted. This scheme is meant to be more balanced and based upon a broader view of the dicipline than each of the individual sources. In order to manage the rather extensive information contained in the description of the existing schemes as well as in the new scheme UCSIE, a database has been designed and used successfully.

UCSIE has resulted iteratively from the various source schemes and their related terminologies (which sometimes differ substantially). Each iteration process specifies a certain level of UCSIE, thereby descending one level. Input is a deliberately chosen set of descriptors in UCSIE at the level to be specified; sometimes the terms used are similar to existing ones, sometimes not. Other required input consists of the collected items of the various source schemes. Allocating these items (quite a number!) to the available UCSIE descriptors yields a comprehensive specification for each descriptor. Conversion of these descriptor specifications into a limited number of new UCSIE descriptors, one level deeper, induces the next iteration. The iterative process then continues.

This procedure would have been hardly possible (and at the best very laborious) to perform without invoking database technology. Also at the end it is quite easy to show the relations between UCSIE and the other schemes.

The comparison of curricula takes place by mapping each individual curriculum table (stored in a database of institutional curricula) onto an UCSIE-based database, while simultaneously assigning the study load related to each subject and a property which is called the 'view' on the subject. By this 'view' it is possible to account for the variety in educational scope for each subject. Four different views can be selected: the view of a conceptualist, of a designer, a practitioner or a generalist. At the end the UCSIE-based database is used to do the statistics throughout the various educational programs.

Results

Preliminary results, using UCSIE, have been obtained for a number of higher vocational institutions. These results show significant differences between educational programs where one actually would expect similarities, while on the other hand similarities occur where differences would have seemed more plausible. This applies for instance to the mathematics part in the curricula.

Conclusion

The main conclusions from the research project at this stage are:

- Database technology allows for a very efficient and detailed comparison of curricula;
- The Unified Classification Scheme for Informatics Education (UCSIE) seems to be an adequate new tool for comparing curricula in a broad spectrum of the informatics discipline;
- Curriculum comparison can create clarity, and simultaneously may raise questions, with respect to presumed and observed differences and similarities. This is highly desirable for both starting students and the consumers of graduated students, the employers.

References

ACM/IEEE-CS Joint Curriculum Task Force (1990). Computing curricula 1991. New York: ACM, and Los Angeles: IEEE-CS.

Coulter, N.S. (ed) (1991). Update to the Computing Reviews Classification System. Computing Reviews, 32(1), 5-50.

DPMA Curriculum Task Force 1990's (1990). Information systems: the DPMA model curriculum for a four year undergraduate degree. DPMA report. Parkridge: DPMA.

Mulder, F., Hacquebard, A.E.N., Veenstra-Strijland, T., & Smeets, D. (1992). Classification of informatics education. Proceedings 12th World Computer Congress IFIP'92, to be published. Amsterdam: North-Holland/IFIP.

Tucker, A.B. (ed) (1991). A summary of the ACM/IEEE-CS joint curriculum task force report: Computing curricula 1991. Commun. ACM, 34(6), 68-84.

THE DEVELOPMENT AND EVALUATION OF A PSYCHOLOGY COURSE FOR ADULT DISTANCE EDUCATION

Welko Tomic, The Open University , Heerlen, The Netherlands

Abstract

The first part of this paper deals with the development of an introductory psychology course for the Dutch Open University.
Use was made of

- a strategy involving systematic design activities and
- the features of a powerful teaching-learning environment. In the second part of the paper a description is given of a formative evaluation of this course. For this purpose a random sample of 400 students completed a questionnaire asking them to evaluate: course content; arrangement of course material; level of difficulty; expectations for the course; opportunities to monitor one's own progress; self-instruction strategies; usefulness of the study guide; use of prior knowledge related to the course; and workload. The responses to most of these questions were positive.

DESIGNING A CURRICULUM EVALUATION FOR A VOCATIONAL PROGRAM IN A LARGE CORRECTIONAL INSTITUTION

Maurice J. Eash, University of Massachusetts, Boston

Introduction

Curriculum evaluations on a large scale present a myriad of problems. Usually randomization is precluded by the scheduling and curriculum requirements of the subjects. The problems of specification of treatment variables and outcomes, as well as obtaining adequate measures and fair comparisons with equivalent populations, plague the curriculum evaluator methodologist. When curriculum evaluations are undertaken in the unusual settings of large correctional institutions, these methodological problems are further compounded by context variables unique to prison settings and disparate population samples. Prison populations have vastly over-representations of populations who: score low on basic achievement tests in mathematics and reading, are deficient on measures of social problem-solving skills, have no proficiency in vocational trade skills, and present scattered, uneven records of unskilled employment prior to incarceration. Social cognitive deficits which parallel these behavioral manifestations have been found to correlate highly with career criminal behavior patterns and subsequently a long history of involvement with the criminal justice system (Rose and Fabiano, 1985). While the consequences of these social cognitive deficits are readily apparent in the behavior of adjudicated youth within the setting of the correctional institution, the measurement and evaluation of curricular and instructional efforts directed to remedy through fundamental change the deficits in the twelve social problem-solving skills common to these youth present unique methodological evaluation research problems. Previous research has identified twelve social problem-solving skills that youthful offenders lack or use ineffectively in their social functioning.: controlling first impulses, taking another's perspective, clarifying and formulating the problem, setting appropriate and realist curriculum organized in the correctional system to address these two major deficits was a modularized, job-specific, competency-based curriculum ordered in scope and sequence to meet the special problem of students who are frequently transferred from facility to facility to continue their education. The curriculum will be implemented over a three-year period of funding and will include 14 vocational areas in ten correctional institutions. An integrated approach in teaching academic skills in reading and mathematics will be used. An evaluation is required during all three

years which will be based upon ". . . acquisition of competencies, the integration of vocational and academic skills and licences or other credentials gained."

Method

An outside evaluator from a state university has been commissioned to assist in designing the curriculum evaluation plan and to conduct the first year's evaluation. The curriculum is being used during the 1991-92 school year and is a slightly modified Massachusetts Competency Based Vocational Education Curriculum which encompasses both the specific occupational areas (N=14) and the mathematical competencies required. Two approaches to evaluation are underway during year one. The curriculum treatment and its effects in year one will be evaluated using experts in the vocational areas being taught who will observe the program, examine the curriculum as taught, and interview the instructors and a sample of students. While they (the experts) will have the major responsibility for judging the curriculum processes and products, they will confer with the outside evaluator on selection of data sources and data gathering procedures.

The second approach to evaluation in year one will involve the teaching faculty in designing a curriculum evaluation which will be implemented in years two and three. Based upon theoretical and empirical results obtained in participant evaluation (Shapiro and Blackwell, 1987), the design will be formulated to assist the teaching faculty, as well as to assess outcomes. A series of workshops will be held with teaching faculty which will involve them in determining effects of treatment, defining and selection of measures, developing comparisons, and analyzing data for specification of treatment effects and drawing of generalizations from findings. The implementation of the curriculum design will be monitored for year one and enter into the evaluation plan for years two and three. The method will be directed to defining the broader specifications of the evaluation and to bringing the curriculum evaluation design into the purview of the teaching faculty's day to day instructional activities. In short, the project tests a paradigm in curriculum evaluation which includes participants in an experimental curriculum in planning and implementing the evaluation as they teach a curriculum new to a specific setting, a correctional institution.

Results

Results will be reported from four data sources:

- the teaching faculty and their reactions to their involvement;
- the students and their reactions to the curriculum;
- the outside experts and their judgments of the processes and products of year one; and
- the quality of the curriculum evaluation design for years two and three.

Some preliminary questions to be examined in the year one study are: what are the constraints on formulating a curriculum evaluation design within the setting of a large correctional institution; and what aids or depresses a teaching faculty's contribution to a curriculum evaluation design?

Conclusion and Discussion

Conclusions and the subsequent discussion will relate the findings to a body of literature on curriculum evaluation methodology, e.g. how best to implement a

curriculum evaluation to contribute to current needs of faculty for information on students and to obtain data on more general treatment effects ascribed to the curriculum. Also, the problem of adoption of a curriculum developed in another setting, public schools, by correctional institutions will be discussed.

References

Larson, K.A. (1988). Remediating problem solving skills. Journal of Correctional Education, 39, 2, pp. 70-74.

Ross, R.R. & Fabiano, E.A. (1985). Time to Think,. Johnson City, Tenn: Institute of the Social Sciences and Arts.

Shapiro, J. & Blackwell, D.L. (1987). Large scale evaluations on a limited budget: the partnership experience, in: The Client Perspective on Evaluation. J. Nowakowski (Ed.). pp. 53-52., Inc. San Francisco: Jossey-Bass.

APPRECIATION OF LIVING ENVIRONMENT AMONG 10 TO 12 YEARS OLD SCHOOL CHILDREN

Leena Aho, Aija Haapala, Ulla Salonen, Joensuu, Finland

Introduction

This poster presents some results obtained in the research project - Man and Environment, An Educational Point of View - carried out in the University of Joensuu. The learning and teaching concerning environment are here examined as a cognitive, socio-emotional, and ethical questions. The connections between cognitive, ethical, and emotional developmental levels have been found to exist in earlier studies.

According to the curriculum of the Finnish comprehensive school environmental education ought to be incorporated in the teaching of every subject, both in natural and social sciences as well as in arts and the humanities. Especially at the primary school level the starting point of environmental education is the pupil's own environment and man's relation to it. Then the attention has been paid to how the environment has been perceived and appreciated by the pupils, and what kind of knowledge and values this appreciation is connected with. In general, natural sciences, especially ecology, its concepts and theories, have provided the scientific basis in analyzing the consequences and the environmental changes in nature caused by man's action. But lately also other sciences have began to play a role in studying man's relation to natural and built environments.

In the aesthetical evaluation of the environment the concepts "aesthetic ecology", "environmental aesthetics", or "ecological aesthetics" have been taken into use. This means that the aesthetical appreciation and experiences depend on ecological knowledge. The environment where man's actions are found to be in harmony with ecological principles is appreciated as beautiful. However, everything that is in accordance with these principles is not necessarily regarded as beautiful.

In this study we try to discover two things: First, what kind of environment do the pupils evaluate as pleasant and unpleasant living environments? Second, what kind of knowledge does appear in the evaluations of pleasant and unpleasant living environments?

Subjects, methods, and material

The subjects of this study were 10 to 12 years old primary school children (n 132) who lived either in a small town or in sparsely populated countryside. Each pupil made two drawings: "In this environment I want to live", and "In this environment I

don't want to live". Before drawing the teacher discussed with pupils what the concept "environment" really means.

Results

The drawings could be classified as follows: The most pleasant living environment was one's own home environment or generally a typical Finnish home environment in a suburban area, (71 %). Foreign environments were presented in 12 % of drawings, a hobby environment in 9 %, a fantasy environment in 5 %, and a luxurious environment in 3 %. In some drawings the social relationship was important.

The most unpleasant environment was a polluted one (54 %), and then came a densely built urban environment, 20 %. An asocial, aggressive, and warlike environment was pictured in 12 % of drawings. A very remote living area (8 %) and other unpleasant environments were also presented.

Particularly the rural pupils presented a home environment with courtyards and with nature elements as the best place to live in, although they did not like the remoteness of the area. The urban children drew foreign, hobby, fantasy, and luxurious environments more often than the children in the countryside.

Discussion

The goal of the curriculum: to appreciate one's own home locality and home land, seems to realize among the children of this age. In the drawings of the pleasant living environment man made environment was portrayed to be in harmony with nature. Further, it could be noticed that they were aware of the ecological consequences: pollution, a topic included in the biology curriculum received more emphasis than asocial and aggressive factors such as violence and war. The problems, caused by individuals, e.g. the pupil him/herself, were not paid attention to in the same way. This result is similar to the way the causes of environmental problems are explained by people of different ages; they are caused by others, e.g. by industry or traffic, but not by the person him-/herself.

The ecological knowledge appeared clearly when the pupils drew the environment where they didn't like to live in. The pictures of the polluted urban environment were numerous compared e.g. with the environment of violence and war. This can mean that the Finnish school children are more afraid of misuse and exploitation of environment than of the other social factors. In the drawings of the older pupils the pollution played a greater role than in those of the younger ones, which could be explained by an increased environmental knowledge and awareness. Because none of the urban pupils presented the town environment as an unpleasant living environment, it can be also claimed that most of the urban pupils appreciated their own environment.

The results confirm the conception that cognitive and value aspects are unseparately involved in environmental education, which in this study means environmental appreciation. Therefore this has to be taken into account in curriculum improvement so that the environment can be examined from different points of view at the same time.

References

Aho, L. (1984). A theoretical framework for research into environmental education. International Review of Education XXX, 183-191.

Aho, L. (1984). Man and nature: Cognitive and emotional elements in the views of twelve-year-old schoolchildren. Scandinavian Journal of Educational Research, 28, 169-186.

Aho, L., Paasi, E. ,& Permikangas, T. (1991). Nature in relation to ethical and aesthetic education. In P. Oja & R. Telama (Eds.), Sport for all. (pp. 589-595) Elsevier Science Publishers B.V.

Sepänmaa, Y. (1986). The beauty of environment. A general model for environmental aesthetics. Annales Academiae Scientiarum Fennicae, ser. B, tom. 234.

Working Group on National Strategy for Environmental Education (1991). A national strategy for environmental education. Helsinki: Publication of the Finnish National Commission for Unesco No 54.

THE IMPACT OF RESEARCH ON THE DEVELOPMENT OF A NATIONAL CURRICULUM: A CAUTIONARY TALE

Wynne Harlen, Scottish Council for Educational Research, Edinburgh, United Kingdom

Introduction

The focus of this paper is the use of research findings in the development of new curricula. It begins with a brief account of the research into pupils' learning in science which, it is contended, should be reflected in current science curricula. Although it may seem to be a diversion, it is necessary to provide a brief overview of the findings of this research since they provide the evidence which testifies to the importance of taking learners' ideas into account in deciding appropriate learning experiences.

A curriculum statement is not neutral with regard to a view of learning and it is not unreasonable to expect a *new* curriculum to take into account the most recent understanding we have of learning. Not to do so deals a blow to educational research. The cautionary tale is of a situation in which the research appeared to be ignored and indeed quite conflicting messages about learning conveyed through the form of the curriculum statements. A teacher's view of learning is one of the strongest influences on his/her choice of learning activities and of the roles which pupils, teachers and materials play in the learning process (Harlen and Osborne, 1985). Thus it is important for evaluation of the implementation of a new curriculum to include the impact on teachers' views of learning.

The tale: research and curriculum change in science education

A gradual but quite fundamental change has been taking place in our view of how children learn science partly as a result of research begun in the 1980s and still continuing. Numerous studies have shown the existence and importance of ideas that pupils construct for themselves about the scientific aspects of the world. At first these were called 'misconceptions' (eg, Johnstone et al, 1977) but soon it was more appropriate and acceptable to call them 'alternative frameworks' (Driver, 1985) or 'children's ideas' (Osborne and Freyberg, 1985) as studies revealed that they were the result of thinking and reasoning by the pupils and indeed made sense to them. Because of this they were difficult to change and certainly were not to be overturned by teaching the 'right' ideas. Some brief examples indicate the nature of these ideas:

- Many pupils at the end of secondary school – and not a few adults – consider that *current from a battery* is 'used up' in a light bulb in a *simple circuit* (Osborne

(1983)). Often the bulb is depicted as receiving the current and thus the complete circuit through the bulb is not necessary. A superficial look at a torch can easily reinforce this notion. It is very much a common sense view that *something* from the battery is used up by the bulb and in many ways it is more reasonable than the scientific view that the current to and from the bulb is the same.

- In science lessons the word *animal* is used to refer to all members of the animal kingdom including human beings but many children use the word more selectively for large land animals, most commonly (Bell and Freyberg, 1985). It is not surprising that some confusion is caused when teachers assume a reasoning based on a scientific meaning whilst the children's understanding is based on everyday usage of words.
- Pupils' views of the relationship between the *movement* of objects and the *forces* acting on them (Gunstone and Watts, 1985) include the ideas that constant motion requires a constant force; that if a body is not moving there is no force acting on it; that if a body is moving there is a force acting on it in the direction of motion
- A very common view of sound held by children is that it travels through air and through holes in solid objects but not through the solids themselves. Also, children often use the word 'vibration' in association with sound but not in the scientifically accepted way; more as cause and effect rather than identity. Again this creates plenty of opportunity for misunderstanding between children and teachers.

The significant features of these examples – and the many other findings of children's ideas – are, first, that they exist and second, that they have an intriguing logic. As adults we can recognise, and sympathise with, the common sense views of the world they represent. We can see that these views emerge from a process of reasoning about experience rather than from childish fantasy or imagination. It is this quality that makes them difficult to replace by the more widely accepted scientific views, which often seem much *less* reasonable in explaining particular instances. The danger of ignoring them in teaching is obvious.

The findings of this research on children's ideas has contributed to new theories of learning; theories developed to explain how pupils form these ideas; theories from which hypotheses can be, and have been, created about how to change them towards the more scientific view. In other words they have proven curriculum relevance. An important question to ask is, did this research have any impact on *the* most important development in science education, the National Curriculum in England and Wales? Several researchers have suggested that it did not and it is hard to refute this by pointing to specific examples. Yet I know from personal involvement that the Working Group members were aware of the research. Some of its members were those who had contributed to the research and members were coopted to add expertise where it was lacking. Knowledge of children's ideas certainly influenced the Group's statements about learning and about the importance of taking children's ideas into account (DES, 1988). But these statements were not part of the official curriculum.

The National Curriculum Working Groups were all set the task of identifying "clear objectives – attainment targets – for the knowledge, skills, understanding and aptitudes which pupils of different abilities and maturity should be expected to have acquired at or near certain ages" (Terms of reference, DES 1988, Appendix A). Moreover it was necessary to identify progression at ten points along a scale to describe the progression in children's learning from age 5 to age 16. The problem

is that it is not acceptable to express as *targets* the 'everyday' ideas we know children to hold at various points in development . For example, even though it is quite normal for infant children of the age of 6 or 7 to consider that growth only happens at night (Russell and Watt, 1990) it would be peculiar, to say the least, to include this as an aim. Consequently the statements of targets in the science curriculum describe a progression in ideas but one which is hierarchical rather than one based on adapting and modifying children's own ideas. There is incompatibility between an objectives model of the curriculum and the developmental model emerging from the research (Harlen, 1990) . Thus the very structure of the curriculum – derived from theories which do not have a research base – seems to preclude the application of the research into children's ideas.
An alternative structure is needed which takes into account research into children's learning - as well as research into how to produce change in the curriculum, which has not mentioned, but which is equally relevant and just as often ignored in national curriculum reforms. An alternative which I believe to be be preferable is to define goals for the end of major stages of schooling - end of primary and end of secondary - and to identify the contributing strands of skills and ideas as dimensions of progress, but to leave the pathways along these directions for the teacher to decide, guided by curriculum materials which take account of how children learn.
It has been argued that, since decisions about the National Curriculum are matters of policy, "researchers' best analytical contribution to policy lies in the measurement of policy outcomes rather than the initial formulation of policy " (a . view deriving from the United States, reported by McGlynn, 1991). Whilst recognising that such decisions must be influenced by other factors as well as research, to reduce research to a role only in evaluating products of political decision would be to endanger the aims of education in democratic society.

Evaluation and the curriculum

However, research does have a role in the evaluation of curriculum innovation. It should enable us to learn from and to correct for, our mistakes by taking a formative role in development. One of the areas of study relevant in the context of the concerns of this paper is the impact of new curriculum programmes on teachers' views of learning. Will the view of learning transmitted through the curriculum influence teachers or will teachers' views of learning established through their training and experience override the messages implicit in the curriculum? To address this research question, *inter alia*, requires an evaluation programme which is able to follow changes in teachers' perceptions and practices as the innovation is introduced. Such a programme is under way in Scotland, where a new central curriculum is being introduced, having a similar structure of targets at successive levels to that of England and Wales (SOED, 1989).
A coordinated programme of four projects looking at different parts of the Scottish 5-14 Development Programme began in September 1991 and will continue at least until 1995. There are two sample of schools: a random sample of about 130 secondary and 200 primary schools who will respond to questionnaires sent in the Autumn for three successive years, starting in 1992; a smaller sample of about 16 secondary and 26 primary schools (linked in feeder-receiver relationships) which will be visited for in-depth interviews with key staff. The first of these rounds of

visits, which generally extend for four days, took place in the first three months of 1992, providing data about actual practices and problems relating to the new curriculum, assessment, testing and reporting. They enable the evaluators to identify issues and generate hypotheses which will be explored with the representative sample through the questionnaire. The intention to return to the same schools for three years in succession provides opportunity to follow up the implementation in a way which one-off evaluations preclude. No analysis of information has yet been undertaken, but it is hoped that it may be possible to share findings at future ECER conferences.

References

Bell, B., & Freyberg, P. (1985). Language in the science classroom In Osborne and Freyberg (eds). (1985). Learning in Science The Implications of Children's Science Heinemann, London.

DES (1988). Science for ages 5 to 16 DES/WO, London.

Driver, R. (1985). The Pupil as Scientist? Open University Press, Milton Keynes.

Gunstone, R., & Watts, M. (1985). Children's Ideas in Science R Driver, E Guesne and A Tiberghien (eds) Open University Press, Milton Keynes.

Harlen, W. (1990). Research and Science Education In Policy Issues and School Science Education E W Jenkins (ed) Centre for Studies in Science and Mathematics Education, University of Leeds.

Johnstone, A.H., MacDonald, J.J., & Webb, G. (1977). Misconceptions in school thermodynamics, Physics Education 12, 248–51.

McGlynn, A.S. Research and policy making in Scottish Education: the role of the Research and Intelligence Unit (RIU), Scottish Educational Review, 23 (2) 112- 117.

Osborne, R.J. (1983). Modifying children's ideas about electric current Research in Science and Technological Education 1, (1) 73–82.

Osborne, R.J., & Freyberg, P. (1985). Learning in Science The Implications of Children's Science Heinemann, London.

Russell, T. & Watt, D. (1990). Growth SPACE Research Report. University of Liverpool Press.

Scottish Office Education Department, (1989). Curriculum and Assessment in Scotland: A Policy for the 90s Paper No 1. SOED, Edinburgh.

COHERENCE AND CONTINUITY 5-16 : ENGLISH LESSONS FOR CURRICULUM MANAGERS

Penelope Weston, National Foundation for Educational Research in England and Wales, United Kingdom

A research study of the implications of the 1988 Education Reform Act on the management of the whole curriculum during compulsory schooling (5-16) in England and Wales identified major questions about the future shape of the curriculum.

Although the Act broke new ground in codifying the content and structure of the National Curriculum (which is likely to occupy the major part of the timetable), it is far from clear that, in practice, it is creating greater coherence across the curriculum as a whole, as it is experienced by pupils and teachers. It has, however, produced a new challenge for schools, in defining Programmes of Study that run from 5-16 with no clear break between primary and secondary phases. Now the problems of providing continuity, especially at points of school transfer, are explicit. The difficulty of ensuring progression for every pupil (given widely disparate rates of individual progress, within and between subjects) is also becoming more apparent.

The study took into account the other, outwardly non-curricular, aspects of the 1988 Act, in particular the local management of schools and 'open enrolment', that is giving more control to schools over pupil recruitment within a 'market' framework. The aim was to consider how these factors, alongside established differences between schools, for example in terms of location, size and innovation history, might be influencing whole curriculum policy. It was also important to review the changing role of the Local Education Authority (LEA) in curriculum policy and support for schools. The study was funded through the National Foundation for Educational Research's Membership Programme, which is planned to meet the concerns of the members, primarily the Local Education Authorities of England and Wales. It was undertaken between November 1989 and April 1991, and endeavoured to provide an evaluative critique which could assist managers and policy-makers in schools and LEAs, while raising issues that would be of interest to a wider educational readership. In order to follow through the impact of the 1988 Act on the curriculum as a whole, case studies were undertaken in groups of schools covering the 5-16 age-range in three LEAs: in all, four secondary schools and 12 primary schools. The three LEAs differed in location, size and approach. These school-based studies were supplemented by regular discussion groups with representatives from these and nearby LEAs, to review emerging issues; and

by a survey of about 700 maintained schools in England and Wales, undertaken late in 1990 to investigate at national level the findings from the case studies.

The starting point for the study was the centrality of the curriculum as the means for promoting learning for all pupils, what Shipman (1990) has called the 'learning package'. In his view, the task for managers is to establish a new equilibrium by integrating this 'learning package' with the other main elements of the system, as it has now been defined following the 1988 Act: the management of resources and the management of those with an influence on the curriculum, which now includes parents and governors as well as central and local government.
The paper will focus on some of the main findings of the study, particularly as they relate to the concept of a coherent, 5-16 whole curriculum framework, incorporating a very large element of prescribed National Curriculum elements. The issues which will be addressed include:

Curriculum cultures in primary and secondary phases, and their significance for whole curriculum design
Since the National Curriculum was defined and designed in terms of subjects, it is difficult to break out of this mould in thinking about the whole curriculum. In secondary schools, with their strongly developed subject cultures, there is little reason to do so. In some schools, the statutory requirements have actually reinforced subject differences and competition between them, and made it more difficult to plan for the curriculum as a whole. In primary schools, however, only the core subjects of maths, language (mother tongue) and some specialist areas such as music had in recent practice been considered, and in some cases timetabled separately in many schools. Instead, much of the curriculum, especially in the early years, was usually organised in 'topics' which covered a number of subject areas. As a result it seems to have come naturally to many primary staff teams to review their curriculum as a whole, when planning how to implement the subject requirements of the National Curriculum. The two phases have therefore adopted different designs and strategies for whole curriculum planning.

Competition or collaboration in developing a 5-16 curriculum?
At a time when each phase has been preoccupied with meeting its own statutory goals, it has been difficult to overcome the traditional barriers to primary-secondary curriculum deliberation. Moreover, the pressure (resulting from the 1988 Act) to recruit the maximum number of pupils has affected relationships between secondary schools, and group of the primaries from which they jointly recruit, particularly in urban areas. Some LEAs have grouped a secondary and its feeder primaries into a family or pyramid structure to discourage rivalry, and also to encourage collaborative curriculum planning and review. Even then, differences of interest, organisation and resourcing between primaries and secondaries often impeded efforts to develop joint curriculum initiatives. The paper will review the constraints, and also the factors which seemed to promote cross-phase collaboration.

Curriculum progression for the pupil

A complex structure of assessment has accompanied the National Curriculum. Will it help or hinder individual progression? In principle, it should be possible to improve diagnostic assessment across the curriculum, using a common framework for recording achievement in all aspects of learning, and use this information to enhance learning for each pupil. In practice, schools were experiencing a number of problems in managing the system, and resented its use as a control mechanism for schools (through the publication of formal test results). The paper will review what some schools were doing to promote pupil progression, across the curriculum, within and across the primary and secondary phases, and consider some major obstacles.

Finally, the paper will discuss the *manageability* of a whole, 5-16 curriculum, in the light of the study's findings. Is the concept realistic, in principle, and can it be realised through the scheme adopted in the 1988 Act? Can the criteria of continuity (of curriculum provision) and progression (for each pupil) be met within that framework?

Note

The study was directed by Penelope Weston and Dr Sheila Stoney, head of the Department of Evaluation and Policy Studies. The other members of the team were James Jamison (Senior Research Officer), Elizabeth Barrett (Research Officer), Neil Rubra (Statistician) and Lynn Fardell and Jocelyn Simmons (Secretaries).

References

Curriculum Council for Wales (1991) The whole curriculum: 5-16 in Wales. Cardiff, Curriculum Council in Wales

National Curriculum Council (1990). The whole curriculum (Curriculum Guidance 3). York: National Curriculum Council

National Curriculum Council (1991). Report on monitoring the implementation of the National Curriculum core subjects, 1989-1990. York: National Curriculum Council

Shipman, M (1990). In search of learning: a new approach to school management. Oxford: Blackwell

CAN THERE BE A TRANS-ATLANTIC CURRICULUM RESEARCH AND/OR CURRICULUM THEORY?

Ian Westbury, University of Illinois at Urbana-Champaign, U.S.A.

Selander (1985) observes that if one sees the ways in which the "territory of pedagogy and education" are written about as analogous to a philosophical or a literary text, then context becomes very important for understanding how pedagogical texts are understood. He goes to suggest that "pedagogical texts are more closely related to context than are literary texts. Pedagogical texts are (mostly) normative and have as a reference, specific actions within defined contexts". As a result, few pedagogical texts successfully cross cultural boundaries, and even fewer cross in a form which reflects their "meaning" in their original context.

Selander's theme is the theme I explore in this paper. My topic is summarized in my title -- but elliptically. What I will be seeking to do in this paper to consider the possibility of an international discourse around the curriculum (see also Westbury 1986). In principle such a discourse should be enriching to the nationally-oriented communities who undertake curriculum research. *But the backdrop against which I will consider this issue is one in which, in the main, curriculum research and theory of the kind I am concerned with does not pass across cultural boundaries.* There is, for example, little or no European research discussed in the recently issued *Handbook of Research on Curriculum* or in the curriculum sections of the latest edition of the *Encyclopedia of Educational Research*. These are of course American-written and -originated volumes and the parochialism of Americans is well known.

But this non-passage of curriculum research is seen also dramatically in papers published in the *Journal of Curriculum Studies* which has, since it beginnings nearly 25 years ago, sought to be an *international* journal, reporting research (principally) from Europe, North American and Australia. However in spite of a long-standing editorial intention to publish papers which speak across cultural regions, only 30 of the 113 citations of the 10 most-cited papers in *JCS* (nine from North America) came from outside the region of origin of the author (and of this 30, six citations were to one paper and 10 of the 30 came from Australia).

Why is curriculum research and theory so parochial? One reason is, of course, institutional, citation and use of papers is part of the on-going interactions of social groups which tend to be nationally- and culturally-based. But, as important -- and this is Selander's claim -- in that it is a specific context which determines the

referents of a given argument, the different contexts of curriculum studies hold very different assumptions about what constitutes a significant argument. I discuss three such contexts which I see as determining the different problematics of U.S. and European curriculum research:

- the dominant ideology and form of curriculum control;
- the extent and forms of resolution of the problem of mass secondary education;
- the form of the structural linkage between the secondary school and university, i.e., the relationship between the secondary school and the "higher culture".

I argue that the difference between the U.S. and Europe in the ways in which these three themes play out has made and makes it difficult , if not impossible, for a trans-Atlantic "school" of curriculum research to emerge. Even in those areas in which an international discourse has emerged (e.g., in the Anglo-American concern for "transformative" ways of thinking about schooling associated with, for example, the work of Michael Apple or the German reception of American "curriculum theory" in the 1970s and 1980s) what is said has, I believe, clearly different meanings in the different contexts of production and reception. Having said this, there are forces which might produce interesting convergences in the contexts of curriculum studies in North America and Europe and which are, in turn, might lead to possible convergence of research traditions. I examine three such areas: *feminist curriculum theory*, research and theory concerned with *the control of the curriculum*, and research on the *teaching and learning of school subjects*.

References

Selander, S. (1985) The pedagogical 'web': An analysis of text related to context: The example of Freire and the Swedish dialog-pedagogy. Paper presented to the annual meeting of the American Educational Research Association, Chicago (Department of Educational Research, Stockholm Institute of Education).

Westbury, I. (1986) Curriculum theory and curriculum research in international perspective. In B.B. Gundem (ed.) Om Læreplanpraksis og Læreplanteori. Kompendium 3 (Pedagogisk Forskningsinstitutt, Universitetet I Olso), 167-200.

COURSEWARE IN THE CURRICULUM: DEMANDS ON TEACHERS

Joke M. Voogt, Department of Education, University of Twente, Enschede, The Netherlands, & John Olson, Faculty of Education, Queen's University, Kingston, Canada

Abstract

Changing demands on teachers make integration of courseware in the curriculum a laborious task. In the three projects presented here strategies are employed and investigated to promote courseware use within the ordinary classroom setting. Carleer describes a teacher–centered implementation strategy and investigated the factors that contributed to a broad and varied integration of computers in the curriculum. Courseware for a process–oriented writing curriculum has been developed by Van der Geest. An extensive formative evaluation was conducted in order to realize well–implementable courseware. Voogt developed courseware for science education. Courseware characteristics intending to influence the planning and coaching behaviour of teachers will be evaluated in her contribution.

Introduction to the symposium

Integration of computer use within the curriculum is a complex undertaking as appears from numbers of studies. Often computer use remains isolated and restricted to activities additional to existing teaching practices. Apparently not only the technical aspects of the computer cause problems, but also the changing demands on teachers make computer integration a laborious task. In this symposium the integration of computer use within the curriculum will be approached from the teacher's point of view.

In the Experimental School Project, schools for secondary education cooperate with a research institute on the theme of computer use in education. The three schools participating in the project are provided with computer and manpower facilities. The project started in 1987 and will continue until 1993. Within the setting of this project, three research studies aimed at the integration of courseware (software and accompanying written materials) in the curriculum.

In the first project Computers in the Pilot Schools a teacher–centered implementation strategy of computers has been employed. In the participating schools a variety of activities for teachers in all curriculum domains were prepared and organized throughout the project . The aim was to investigate which factors contribute to a broad and varied integration of computers in the curriculum.

In the two other projects implementation processes within a specific curriculum domain were studied. In the Computer–assisted Writing Instruction Project and the Computerassisted Lab Work Project courseware has been designed from an implementation point of view. The aim of both projects was not only to integrate the computer in classroom instruction within the language and science domain but also to achieve innovative curriculum goals. The Computer–assisted Writing Instruction Project focused on a process–oriented approach to written composition. The Computer–assisted Lab Work Project concentrated particularly on students' inquiry skills. So it can be said that in both projects the teacher who is supposed to use the courseware in classroom instruction has to deal with two innovations at a time. In the Computer–assisted Writing Instruction Project teacher–related factors affecting the implementation of courseware were identified. In the Computer–assisted Lab Work Project courseware characteristics intending to influence the planning and coaching behaviour of teachers were evaluated.

A STRATEGY FOR THE INTEGRATION OF COURSEWARE BY TEACHERS

G.J. Carleer, Department of Education, University of Twente, Enschede, the Netherlands

An aim of the project "Computers in the Pilot Schools" is to investigate a strategy for the integration of computers and courseware in the curriculum by teachers. This project is carried out at two pilot schools in collaboration with researchers of the University of Twente. From the first beginning of the project the focus was on questions and problems of the school as a whole. A case study approach has been carried out in order to take the context into consideration. An important task charged to the researchers is to make the results of their investigation available and usable for other schools.
The stategy for the integration of computers and courseware proposed by the researchers and carried out by the schools can be characterised as a broad multiplier oriented strategy. "Multipler" means that some teachers who have already somewhat experience in computer use or are interested in computer use monitor their collegues. Their most important tasks are to diffuse the information they received from the researchers and from the computer coordinator, and to support the use of computers by their collegues. This reflects the intention to make use of the within school expertise. "Broad" means that each subject area has his "multiplier teacher".
The research is focussed on the results, the problems and advantages of the multiplier strategy.
Data have been colected in different ways. E.g. four times a state–of–the–art questionnaire about computer use has been distributed among all the teachers of the pilot school at the beginning of each school year; lessons, discussions and in–service sessions of teachres have been observed; teacher repports have been analyzed.
Some conclusions of the study are that a multiplier strategy is effective–although at a very low pace– and avoid a great distance between the computer minded

forerunners and other teachers. Computers have beeb used in the the compputer lab as well as in the classroom. A real integration in the curriculum is still an exception. Most of the computer use is additional.

DESIGNING AND TESTING FOR IMPLEMENTATION

Thea M. van der Geest, Section of Applied Linguistics, University of Twente, Enschede, the Netherlands

The project Computer–assisted Writing Instruction (1985–1990) was initiated to investigate claims that the computer could serve as an useful tool in process–oriented writing instruction. As appears from survey studies of writing achievements of students of secondary school level, most students experience severe problems in producing texts that fulfil their communicative function adequately. This could be due to the little attention that usually is paid to other aspects of writing skill than the 'mechanicals' of text production. In the project an attempt was made to develop courseware for an innovative type of writing instruction in which much attention is paid to defining the communicative goal of the text to be produced and to planning the text with an open eye for the needs of the reader and the consequences for the content of the text.
The material developed should support student writers of secondary school level in their processes of text planning and text production. It was to be used in the traditionally non–machine–supported language arts courses. But especially it was meant to be feasible both for teachers and students in ordinary classroom situations.
To achieve such innovative, well–implementable courseware, a cycle of formative evaluation activities was planned, in order to assess problems teachers and students actually met when using the materials. As soon as plans for or parts of the new materials were available for testing, they were commented on and tried out by users of the target group, that is both teachers and students. A variety of instruments has been used, chosen according to the evaluation question at hand. In an extended field test with eight classes of the target group of students, data were collected and analyzed with regard to factors known to be influential from the educational innovation literature. In a cyclic process, the experimental materials were repeatedly revised on the basis of the results of the evaluation activities. Subsequently the feasibility and effectivity of the revised materials has been assessed in an effect study in nine classes of secondary education (N = 200).
In the paper presentation, attention will be paid to the design of the formative evaluation study and the revision based on its results. The methods used will be informally evaluated for their merits and demerits with regard to usability and practicality. The process of revision on the basis of the data collected will be described.
The following study of the feasibility and effectivity of the revised materials, focused at students' and teachers' perception of the materials and the changes in achievements that could be related to the use of the materials. Quantitative and qualitative measures were combined. Results with regard to the feasibility of the writing courseware, particularly as experienced by the teacher will be reported.

Teacher perceptions of the characteristics of the materials and their use appeared not to be completely parallel to the students'
perceptions and their writing achievements, measured in the effect study.

COURSEWARE CHARACTERISTICS AND THE TEACHER

Joke M. Voogt, Department of Education, University of Twente, Enschede, the Netherlands

In the 'Computer–assisted lab work project' the computer has been used as a tool for collecting and displaying data during students' lab work. The courseware, a comprehensive set of laboratory exercises on 'Heat and Temperature', a student–textbook and a teacher guide, intended to improve students' inquiry skills, particularly interpreting graphs and drawing conclusions. The courseware promoted hands–on learning experiences. It was supposed that most teachers participating in the project were not used to hands–on learning. So the teacher guide payed special attention to recommendations for coaching students while learning by experience.
Goodlad et al. (1979) proposed a typology for different manifestations of the curriculum: ideal, formal, perceived, operational and experiential. In the symposium characteristics of the courseware (the formal curriculum) which positively or negatively influence the desired planning and coaching behaviour of teachers (considered critical in the operational curriculum) will be highlighted.
To describe the implementation of the courseware a case study approach has been used. The cases in the study were five teachers, teaching in six lower secondary science classes (two schools, three classes per school). All lessons (N = about 60) with the courseware were audiotaped and observed, teacher–student interactions could therefore be carefully registered. A structured interview with the teachers about their perception of the courseware goals, about lesson preparation and execution and about effects on students, took place after all lessons were executed.
To analyze the observed lessons a so called 'curriculum profile' (Van den Akker, 1988) has been developed and validated by experts. In the curriculum profile the essential lesson elements were described in operational terms, with a differentiation between 'threshold', 'ideal' and 'unacceptable' elements. Information about the exact amount of time spent to each lesson and about technical problems with hard– and/or software could be derived from the lesson observations.
At this moment data–analysis is still in progress thus the final results are not yet available. Results about the following issues can be expected: (a) which characteristics of the courseware promote or inhibit the desired teaching and coaching behaviour; (b) do teachers change their teaching and coaching behaviour during the execution of the courseware; (c) for what reasons teachers show or don't show the desired teaching behaviour; (d) if and in what way side effects threaten the implementation process.
The findings of the study will be related to literature about implementation and courseware characteristics.

References

Akker, J. J. H. van den (1988). Ontwerp en implementatie van natuuronderwijs [Design and implementation of science education]. Lisse: Swets en Zeitlinger.

Goodlad, J.I., Klein, F. & K.A. Tye (1979). The domains of curriculum and their study. In: Goodlad,J.I.(ed), Curriculum inquiry. New York: McGraw Hill Book.

SEARCHING FOR MORE EFFECTIVE CURRICULUM DEVELOPMENT STRATEGIES

Jan van den Akker, University of Twente, Department of Education, Enschede, The Netherlands

Introduction

Since 1988 we have been engaged in collaborative research activities aimed at increasing the effectiveness of professional curriculum development in our country (The Netherlands). Several stages in this endeavour can be discerned:

- A preliminary study (problem analysis; conceptualization) [1988–1989].
- Retrospective case studies of 18 completed development projects [1989–1990].
- Devising promising procedural guidelines for the design, construction, evaluation, and implementation of curriculum plans and materials [1990–1992].
- Systematic, empirical validation and refinement of these guidelines [1992–?].

This symposium will deal with the rationale for and the preliminary outcomes of this research in progress.

Perspective and framework for the research

The organization of curriculum development in the Netherlands reflects the general educational policy of balancing between central regulations and decentral autonomy. The National Institute for Curriculum Development (SLO) publishes curricular plans and materials that have only an exemplary, not an obligatory, status. Everyone is free to use these products and to adapt them to their own situation and needs. This agreement has been made for political reasons (constitutional freedom of education), but also because of economical (relation with educational publishers) and innovative (stimulation of active use) arguments. In terms of Walker (1979, 1990), one can characterize the organization as aiming at interaction between 'generic' curriculum development by a center (SLO) and 'site specific' activities by schools and teachers, while also various 'intermediairy' organizations (e.g. educational publishers, inservice training institutes) play an important role.

The SLO, in operation since 1976, has carried out dozens of curriculum development projects for (subjects in) primary, secondary, vocational, and adult education. In recent years there was a growing concern about the lack of reliable and systematic knowledge of the development strategies that actually have been used in the past. This was seen as an obstacle to learn from previous experiences and to improve the professional methods, procedures, and techniques for curriculum development. The need for this improvement was strengthened by the

202

results of several evaluation studies on the impact of SLO projects that raised serious doubts about the effectiveness of the development activities.
These concerns have led to this series of studies, starting with a preliminary analysis of SLO policies and practices and an extensive study of the international literature on curriculum development strategies. This resulted in a first report (van den Akker, Boersma & Nies, 1989) that included a conceptual framework for the studies. Authors like Connelly (1978) and Schaffarzick (1975), discussing the generalization problems of research (usually case studies) on curriculum development strategies, urged the formulation of such a conceptual framework. This framework (specifying activities within several core components of curriculum development like: orientation and foundation; design and construction; evaluation and revision; diffusion and implementation) has guided the data collection and analysis in the case studies. It is also useful in the current efforts to devise and test more effective strategies for components of the development work.

The retrospective case studies
The main question for the retrospective case studies was:
What development strategies have been practised by SLO projects?
From the completed SLO projects a representative sample of 18 projects, spread over different educational sectors and subjects, has been drawn. Each of these projects was studied along the following procedure: (a) introductory talk with the (former) project director; (b) document analysis (project publications, curriculum products, evaluation reports); (c) individual interviews with the project director and another member of the development team, using an extensive checklist that was based on the conceptual framework; (d) writing of a preliminary case report; (e) comments from the respondents
and co–researchers; (f) writing of final case report.
Three separate volumes with case descriptions were produced, while in a single book (van den Akker, Boersma & Nies, 1990) the results of all cases were analyzed. The outcomes have been discussed at a symposium with a large group of SLO developers, leading to recommendations and plans for future training and research activities (Keursten, 1991).

The follow–up studies
Next the emphasis of the research has been shifted from retrospective analysis to the devising of procedures that may guide curriculum developers to (more) effective ways of action. The primary focus is on core tasks like the design, construction, and formative evaluation of curricula. Although interrelated, a distinction is made between the development of different kinds of curriculum products like: documents with attainment targets; curricular frameworks at a national or school level; exemplary instructional materials.
The procedural guidelines are devised on the basis of: reflections on the previous case studies findings; additional and focused literature study; analysis of the approaches in the current development practices; efforts to translate relevant research findings in selected areas. The guidelines are formulated in dialogue between the researchers and several small groups of SLO developers in order to further the relevance, acceptability, and feasibility of the proposals for the curriculum development practitioners. Parts of the guidelines are already tried on a

small scale in several contexts; more systematic testing will start during the year 1992.
The ultimate aspiration of this long term study is to contribute to the formulation of insightful, empirically validated, and practical answers to two central questions proposed by Schaffarzick (1975): (i) What procedures are essential for 'high–quality' curriculum development?, and (ii) What are the most efficient ways to use these procedures?

The different presentations
The first presentation (by *Anca Nies*) will describe the results of the case studies. Patterns of the practised strategies in the different core components of the development activities will be analyzed and illustrated. Also more general perceptions of the (rather informal and not systematic) nature of the development practice will be discussed and confronted with the process–oriented 'quality' criteria for curriculum development as described by authors like Short (1983) and Walker (1979).

In the pursuit of quality improvement the second presentation (by *Kerst Boersma*) will describe the efforts to improve the effectiveness and efficiency of SLO projects through a standardization of the decision making process in the planning and execution of development activities. Guidelines for more systematic project management will be presented, including more specific product requirements and development routes.

The third presentation (by *Gert van den Brink*) will focus on procedural guidelines for the development of attainment targets and their possible meaning for core curricula. These guidelines have a firm base in practical development experiences in recent years in several projects for junior secondary education.

The fourth presentation (by *Jan van den Akker*) will offer procedural guidelines for the development of the curriculum materials for the micro–level of teaching and learning. The proposed approach reflects a kind of developmental research with a lot of attention to formative evaluation activities that anticipate on implementation issues.

References

Akker, J. van den, Boersma, K., & Nies, A. (1989). Development strategies in the SLO practice: Report of a preliminary study [in Dutch]. Enschede: SLO.
Akker, J. van den, Boersma, K., & Nies, A. (1990). Development strategies in SLO projects [in Dutch]. Enschede: SLO.
Connelly, F. (1978). How shall we publish case studies of curriculum development. Curriculum Inquiry, 8, 73–82.
Keursten, P. (Ed.) (1991). Report of the SLO symposium "Strategies of curriculum development" [in Dutch]. Enschede: SLO.
Schaffarzick, J. (1975). Questions and requirements for the comparative study of curriculum development procedures. In J. Schaffarzick & D. Hampson (Eds.), Strategies for curriculum development (pp. 211–256). Berkeley, CA: McCutchan.

203

Short, E. (1983). The forms and use of alternative curriculum development strategies: Policy implications. Curriculum Inquiry, 13, 43–64.
Walker, D. (1979). Approaches to curriculum development. In J. Schaffarzick & G. Sykes (Eds.), Values conflicts and curriculum issues: Lessons from research and experience (pp. 263–290). Berkeley, CA: McCutchan.
Walker, D. (1990). Fundamentals of curriculum. San Diego: Harcourt Brace Jovanovich.

SOME RESULTS FROM IEA'S STUDY 'COMPUTERS IN EDUCATION'

Tjeerd Plomp, Department of Education, University of Twente, Enschede, The Netherlands

Introduction

Many countries have adopted policies for the systematic introduction of computers in education. The major goal of the Computers in Education study (Comped) of the International Association for the Evaluation of Educational Achievement (IEA) is to collect longitudinal and cross-national comparative data in order to contribute to the evaluation of policies on (the introduction of) computers in the countries that are participating in the project and to building a knowledge base from which answers to questions about what and how to use computers in education can be sought.

This symposium will focuss on who participated in this study and on the results in a few countries.

Goals and design of the study

The major goals of the study are to describe and analyze crossnationally as well as longitudinally how computers are used in schools by teachers and students, and what cognition, skills and attitudes students have with respect to new information technologies.

The study consists of two stages. During stage 1 (1987-1990) data were collected in elementary, lower secondary and upper secondary schools at school and teacher level. In stage 2 (1991-1994) measures from stage 1 will be repeated and, in addition measures at student level will be taken.

The measures taken in stage 1 of the study were based on a conceptual framework characterizing the educational system in terms of levels of decision-making and identifying the factors contributing to effect changes. These factors were taken from literature on educational change (e.g.: Fullan, Miles, & Anderson, 1988) such as the quality, clarity and relevance of the objectives and the characteristics of the innovation (content, materials, instructional strategies); support and leadership; staff development; experiences with innovations; and the existence of evaluation and feedback. The framework reflects the hierarchical structure of most educational systems, but acknowledges that decisions which promote or inhibit the implementation of computer-related curricula are made at all levels, which may cause discrepancies between decisions and expectations that

206

exist at different system levels. An identification of these discrepancies may in itself be an important starting point for improvement measures in education.
Three populations were defined. Population I are the final grades of elementary education. Population II is lower secondary education and Population III is upper secondary education. Representative stratified random samples of schools and teachers in schools were drawn for each country. A distinction was made between schools using and not using computers.
In stage 1, altogether, by means of questionnaires data were collected from about 70.000 respondents (principals, computer coordinators and teachers) from schools sampled in 21 educational systems.

Countries participating in stage 1
The following countries (educational systems) are participating in the study: Austria, Belgium (Flemish),Belgium (French), China, Israel, Italy, Canada (British Columbia), Japan, Luxembourg, France, F.R. Germany, Greece, Hungary, India, Netherlands, New Zealand, Poland, Portugal, Spain, Switzerland, USA.

Instruments for stage 1
National policy data were collected with a questionnaire which addressed issues like national policies, for example with respect to hardware provision, courseware development, teacher training, budgets, innovation strategy, etc.
A Principal and a Computer Coordinator Questionnaire addressed issues like school policies in using computers, availability and acquisition of hard- and software, organization of computer use on school level, support, equity, attitudes and school characteristics.
Questionnaires for teachers of Computer Education, Mathematics, Science and Mother Tongue contained questions about computer education, types of computer use, frequency of use, time spenditure, curriculum content covered, attitudes, teacher knowledge and skills, and teacher training.

Instruments for stage 2
In stage 2, the same instruments will be used as in stage 1. In addition instruments have been developed to measure functional information technology abilities, attitudes and computer experiences of students. The pilot testing of these instruments took place in 1991 in 10 countries, whereas the main run data collection will take place in Spring 1992.

POLICY IMPLICATIONS FROM THE IEA SURVEY COMPUTERS IN EDUCATION

Willem J. Pelgrum and Tjeerd Plomp, University of Twente-OCTO, Enschede, The Netherlands

In order to determine which information needed to be collected in this study, a framework was developed which identified the key factors at which the study was aimed. The framework consists of concepts derived from systems theory,

curriculum theory and theories on educational change, as discussed in the previous section.

An educational system is a complex of subsystems at different levels: at the macro level the educational system of a country or state, at the meso level the school, and at the micro level the classroom and the student. On each level, educational decisions are influenced by different actors; for example, at the school level the school board, the principal, the subject matter department, and the teacher. External influences may be exerted by, for example, business and industry, or parents. The output of a subsystem at a certain level can be conceived as the input for the subsystem on the next level. For example, the output at the macro level may consist of policies, intentions and plans of governments, laid down in official documents, or existing as shared conceptions of what is expected from schools. Conceiving this as the input for schools, the output at this level consists of the activities and the practices in the classrooms, the time allocations and the instructional practices with computers of teachers. This is the input at the micro level, resulting in activities, cognitive skills and attitudes of students.

In curriculum theories, a distinction is made between the intended, implemented and attained curriculum. The intended curriculum refers to the curriculum plans (at the macro level), which may be laid down in official documents or which may exist as shared conceptions of what the important curriculum content is. The implemented curriculum (at the meso level) consists of the content, time allocations, instructional strategies, etc. which the teacher is actually realizing in his/her lessons. The attained curriculum (at the micro level) is defined as the cognitive skills and attitudes of students as a result of teaching and learning.

Taking these three curriculum levels as major input/output categories one may wonder how these levels influence each other and which factors may explain the occurrence of discrepancies. The literature on educational change may be used for tracing potential factors (e.g. Fullan, Miles, & Anderson, 1988). These factors include the quality, clarity and relevance of the objectives and the characteristics of the innovation (content, materials, instructional strategies); support and leadership; staff development; experiences with innovations; and the existence of evaluation and feedback.

This study incorporates the three different perspectives which are described above. The global conceptual framework for the study, in which the three perspectives are related to each other, will be presented in the article as an introduction to the other articles. Moreover, a description of results on the key factors will be provided. Furthermore the policy implications resulting from the results presented in the other articles in this symposium will be summarized.

WHAT FACTORS INFLUENCE THE RATE OF USE OF COMPUTER USE IN SECONDARY EDUCATION?

Karl Frey, Ruedi Niederer, Institute for Behavioral Sciences, Swiss Federal Institute of Technology Zurich, CH

The given situation

In Switzerland, the computer is firmly established throughout the secondary system. At present approx. 75% of students in compulsory secondary schooling, and 100% of those at the upper secondary level (16+) receive an introduction to computer science. This introduction is given within the separate subject *Informatik*. In other subjects, computers play only a marginal role. One class in seven uses the medium in learning another subject in the course of a year. Translating this result into practical school terms, this means that every few weeks or every few months the teacher performs a demonstration on the computer. It is an even rarer occurrence for the students to work at microcomputers themselves.

Thus we are running the risk that our schools will follow their usual pattern of organization and relegate the computer to its own isolated cell. This could lead to problems of motivation and efficiency, because in school, unlike in the outside world, the computer operates in isolation.

For this reason it is a prime task to explore the reasons for this low rate if we wish to achieve an appreciable increase in the use of computers in school subjects other than "straight" computer science.

Influencing factors

There are a large number of factors which influence the rate of use. First and foremost, the machines themselves.

- If Macintosh computers are available, the level of use is one third higher than in the case of MS-DOS. Correspondingly, significantly more teachers report difficulties in using software if MS-DOS computers are in use rather than Macintosh.
- A further factor which affects the rate of use is the type of school. The level is highest at the *Gymnasium* (terminating at grade 13 with the general university entrance qualification), whilst surprisingly it is the commercial vocational schools, whose students are familiar with computers from their office experience, where the rate of use is lowest.
- The subject also plays an important role. Mathematics classes have twice the level of use of first language classes, although computers as word processors are well suited for use in the whole spectrum of language education.
- Location is another important point. If the computers are in the classroom, the rate of use almost doubles in comparison with situations where the computers are located only in a specially designated computer room. Accordingly, the availability of portable or movable equipment proves an advantage in increasing use.
- Finally, the teachers themselves have a decisive effect on the rate of use. Surprisingly, the age of the teachers has no influence. Young teachers are no more likely to use computers than their older colleagues. On the other hand, gender proved, as anticipated, to affect the rate of use. Male teachers are more likely to use the computer in teaching their subject than female teachers.

- Training also plays an important role. Teachers who use computers in their teaching attend more than twice as many hours of inservice training than teachers who do not use computers.
- The most significant factor here would appear to be the presence of computers in the teachers' homes. In the case of teachers who have their own computer at home, the rate of use is three times higher than in the case of teachers with no computer.

To sum up it may be concluded that the rate of use of computers in the teaching of all school subject depends on a large number of factors. It is important to note that none of these influences are dominant, i.e. it does not suffice if just one factor is positively influenced for the rate of use to be boosted. A range of factors must be positively influenced.

GENDER AND COMPUTER USE IN SCHOOL WORLDWIDE

Georgia Kontogiannopoulou-Polydorides, Department of Education, University of Patras, Greece

The issue of the differentiated use of and access to educational technology by gender has received widespread attention in research in recent years. Computer use and access particularly has been a major concern of researchers and policy makers and is one of the question of the worldwide research project Computers in Education from the International Association for the Evaluation of Educational Achievement (IEA).

The issue of gender equality in education, and especially with respect to science and math related tracks of schooling is a major concern of educational policy analysts since it reflects:

- access patterns of boys and girls to university fields of study in science and technology;
- future division of labor by gender and especially in science and technology related occupations.

Educational policy has reflected these issues more or less in specific countries depending on the degree of social awareness and the influence of social prevailing movements.

The issue of gender in computer education raises once again these concerns from a new perspective:

Computers are increasingly becoming very important in the production of knowledge and at the same time an integral part of the means of production in all spheres of economic activity. The concerns, therefore, which have preoccupied educators worldwide with respect to gender in relation to science and math, are expanding to include the computer.

Thomas (1987) is concerned about cultural differences in the acceptance of educational technology, where "cultural" means the "material culture and the patterns of thought and social interaction that typify a group of people". Others indicate that there is a difference in the frequency of use of the computers in school

by gender (Collis et. al. 1989; Jensen and Klewe, 1989) or in the way person to person to machine interaction is pursued (Siann, 1990; Pozzi, 1992).
In other cases gender differences in computer related studies are masked by bureaucratic routine responses indicating equality where it might not exist (Whyte, 1986).
Still other studies report that gender has little or no relationship to computer experience or computer use (Mc Coy and Dodl, 1989); or computer literacy and achievement for student teachers (Woodrow 1991a; Woodrow 1991b).
The first phase of the Computers in Education research project dealt with the issue of gender as context characteristic of schools using or not using the computer.
The underlying question in this case is whether computer use and education in the countries and schools involved is related to specific context in which gender specific leadership, teaching force and policies are prevailing.
The results show that there is no specific pattern, at the moment and, in the countries participating in the study, in the relationship of the overall schooling context and computer use with respect to gender. That is, there is no evidence across countries that schools with female dominance in their teaching staff tend to use less the computer. On the other hand there are some countries in which this seems to be the case.
But overall it is clear that female dominance in the teaching staff changes more across educational levels (primary, lower sec., upper sec.) rather than from computer using to non-using schools at the same educational level.
On the other hand the overall picture indicates that the pattern of policy adoption is quite similar across educational levels and it does not show considerable changes when examining the total number of schools and the schools with male principals.
Finally concerning the attitudes expressed by principals of computer using schools, women (who agree with a specific statement) tend to spread in the range of percentile points across countries much more that men do, which indicates that men educators tend to accept a uniform educational ideology with respect to computers, while women educators do not. Or that men educators have reached some kind of cohesive consensus regarding the importance of computers in educational and teaching activities, in learning outcomes and teacher training. Women educators have not reached such a consensus across countries.

References

Collis, B. A., Hass, H. and Kieren, T.E. (1989). National Trends in Computer Use Among Canadian Secondary School Students. Jounal of Research on Computing in Education, 22, Fall, No 1, pp. 77-89.

Jensen, P.E. and Klewe, L. (1989). Gender Differences and Computer Use in Education. Copenhagen: The Danish Institute for Educational Research.

Mc Coy, L.P. and Dodl, N. R. (1989). Computer Programming Experience and Mathematical Problem Solving. Journal of Research on Computing in Education, 22, No 1, pp. 14-25.

Pelgrum W.J. and Plomp, Tj. (1991). The Use of Computers in Education Worldwide. Oxford: Pergamon.

Pozzi, S. et al. (1992). Towards a Methodology for Analysing Collaboration and Learning in Computer-bazed Groupwork. Computers in Education, 18. No 1-3, 00. 223-229.

Siann, G. et al. (1990) .The Effect of Computer Use on Gender Differences in Attitudes to Computers. Computers and Education, 14, No 2, pp. 183-191.
Thomas, R.M. (1987). The Nature of Educational Technology. In: Thomas, R. Murray and Victor N. Kobayashi. Educational Technology. Oxford: Pergamon, Chapter 1.
Whyte J. (1986). Girls into Science and Technology. London: Routledge and Kegan Paul, Chapter 11.
Woodrow, J.E.J. (1992a). Locus of Control and Computer Attitudes as Determinants of the Computer Literacy of Student Teachers. Computers and Education, 16, No 3, pp. 237-245.
Woodrow, J.E.J. (1991b). Determinants of Student Teacher Computer Literacy Achievement. Computers and Education, 16, No 3, pp. 247-256.

IMPLEMENTATION OF COMPUTERS IN SCIENCE EDUCATION

Manfred Lang, IPN, Kiel,Germany

The integration of information technology in different school subjects is a demand, which should also be realized for biology, chemistry and physics education. Different studies in the past show, that computer use in science subjects is limited to few topics and occasions, especially for physics (Labudde, 1989; Fauser et al., 1990; Harreis et al., 1991; v.d.Akker, 1991).
Whithin the study 'Computers in Education' of the IEA (International Associa-tion for the Evaluation of Educational Achievement) questions about computer use in science teaching were added to the international standard questionnaires. These questions are supposed to identify difficulties and opportunities to use computers and software in science subjects.
In 1989 questionnaires were sent to a sample of teachers in 600 secondary level I and 399 secondary level II schools. About one half of the teachers selected were using computers in classes, the other half did not.
1123 questionnaires from teachers with different subjects were sent back for analysis. From these teachers 57 of all the 219 science teachers in the study were teaching biology, 61 chemistry and 101 physics.
Whithin the sample 135 science teachers were using computers in classes either for demonstration or for students work. 84 science teachers did not use computers in classes, although 50% of them use a computer at home and 23% of them indicate a high interest to use a computer in class.
Computer use for demonstration by the teacher is more frequent in science classes on secondary level II than in other classes for basic or advanced information technology or maths. On secondary level II 40 out of 65 (or 62%) of the computer using science teachers use computers for demonstration. This percentage is rather high compared to maths (47%) and advanced information technology classes on secondary level II (5%) and basic information technology classes on secondary level I (13%).
No computer use or only computer use for demonstration may in some cases depend on didactical reasons, which do not allow students to handle computers in

science courses. But it appears from the analysis of different questions that the inavailability of computers for the specific purposes of science learning and other problems prevent computer use by the students:

In general science teachers indicate a positive computer readiness. This is measured by a scale composed of eight items about attitudes of computer use in education, interest, reading about and use of computers outside school. The readiness scale is ranging from 1 to 9 with a calculated midpoint of 5.1. The mean scale value of science teachers is 5.9. Expressed in percentages this means that the majority of 62% of the science teachers have positive scale values of readiness. 23% are indifferent and only 15% are sceptical.

In spite of science teachers interest and readiness, computer use in classes is still limited to few topics and by different problems. Only 21% of all computer using teachers are teaching sciences, especially physics. The majority of teachers are using computers in informatics and mathematics courses. The frequency of computer use in science teaching is low: only in some weeks during the school year. In informatics courses for comparison computers are used every week.

Topics with computer use on secondary level II are sound, time/distance/ velocity, waves, optics and nuclear theory for physics and ecology, genetics and evolution for biology. On secondary level I teachers mentioned no specific topics for science teaching. Altogether there are 35 topics mentioned by all 135 computer using science teachers. This number is small compared with 220 topics from 175 maths and 733 from 163 informatics teachers.

In science teaching computers are especially used to measure, control and analyze data and simulate. These kinds of application are most often used in physics and occasionally in chemistry and biology teaching. On secondary level II simulation is used more often than the other applications.

Problems of hardware, software, lesson planning, organisation and others are analyzed with a list of 44 statements. On the average teachers for biology stated 10.0, for chemistry 9.0 and for physics 9.2 problems. The overall mean of all teachers problems in 9.1. Biology teachers mention significantly more problems than teachers with other subjects except informatics. Most often biology teachers complain about not enough time sceduled for computer use, the insufficient number of computers, no adequate room to locate computers and no time to prepare lessons.

In general teachers complain about lack of time sceduled for computer use (49%). A special problem for science teachers is the location of computers, which is most often a computer-room and only for 7% the special science classrooms. Main reasons not to use computers in science teaching are lack of knowledge and skills and insufficient information and guidelines for lesson planning but also insufficient information about software and poor quality of handbooks and materials.

As a consequence from these results it might be possible to develop further computer use in science teaching supported by a high degree of computer readiness. But teachers should be given more opportunities of training, especially to raise their knowledge about basic concepts, applications and software. In addition teachers should have more time for lesson preparation with computers and computers should be available to a higher extent in science classrooms (as IEA-data show for the USA in Pelgrum/Plomp, 1991).

References

Labudde, P.(1989). Computer im mathematisch-naturwissenschaftlichen Unterricht. Forschungsergebnisse aus den USA. MNU, 42/4.

Fauser, R., & Schreiber, N. (1989). Jugendliche, Computer und Bildung: Ergebnisse einer empirischen Untersuchung bei Jugendlichen in 8. Klassen und deren Eltern. BMBW, Studien zu Bildung und Wissenschaft 77. Bad Honnef.

Harreis, H., Riede, H.-M., & Treinz, N.(1991). Ergebnisse einer Umfrage zum Computereinsatz im Unterricht unter SchÅler(inne)n und LehrkrNften in NRW. PhuD 1.

van den Akker, J., Keursten, P., & Plomp,T. (1991). The Integration of Computer Use in Education. International Journal of Educational Research.

Pelgrum, W.J., & Plomp, T. (1988). The IEA-Study 'Computers in Education': A Multinational Longitudinal Assessment. In: Lovis, P., Tagg, E.D. (Eds.): Computers in Education. Elsevier, North-Holland.

ATTITUDES OF SCHOOL PRINCIPALS AND TEACHERS TOWARDS COMPUTERS

Willem J. Pelgrum, Centre for Applied Research on Education (OCTO), University of Twente, Enschede, The Netherlands

The study collected attitude measures on three scales (Educational Impact, Social Impact and Training Needs) from principals in schools that use computers as well as schools not using computers. Moreover, the same scales were presented to teachers. Also measures of teacher's knowledge and skills in using computers were taken.

This paper contains a description of these attitude measures and explanatory variables (like number of years of experience, the innovation context, perceived effects, etc.).

STAFF DEVELOPMENT AS CONDITION FOR COMPUTER INTEGRATION

Ingeborg Janssen Reinen, & Tjeerd Plomp, Department of Education, University of Twente, Enschede, The Netherlands

Introduction

Staff development is a basic and necessary component of the continuing preparation of teachers, administrators and other staff as they extend their professional or technical knowledge (Orlich, 1989). Codianni and Wilbur (1983) compared the findings of seventeen major studies on effective schools and found systematic staff development as being one of six important aspects of these effective schools. Besides normal updating of knowledge in one's subject, professional development and assistance are important for both the dissemination

and implementation of educational innovations (Fullan, 1991). Teachers need to learn new roles in order to work effectively with new programs and technologies (van den Akker, 1988). It is therefore no surprise that staff development and training is one of the factors that determine the success of implementing an innovation in educational practice (Fullan, 1991). There are no reasons to expect that this should not also hold for the use of computers in education. The crucial role of staff development in the introduction of computers in the school is also stressed by Moskowitz and Birman (1985), Walker (1986), Brody (1987) and the US Congres (1988).

Since the introduction of the computer in education, many authors claim that the beneficial aspects of computer use in education are related to the integration of computers in the existing subjects (e.g. Collis, 1988, Hunter, 1984). However, the findings of the first stage of the International Association for the Evaluation of Educational Achievement (IEA)-international research project 'Computers in Education' (Comped) revealed that the use of computers in existing subjects as aid in teaching and learning is limited. In most countries participating in this study introduction of the computer in schools is predominantly happening through the introduction of a new subject like computer education or informatics.

One may argue that computer use in existing subjects is not fully integrated because this form of using computers is the most complex part of the innovation. As Tobin (1987) states, "... adequate hardware, high-quality courseware and administrative support are important, but the teachers' attitude and consequent behaviour will determine the impact of the microcomputer in the classroom". This leads to the conclusion that the limited use of computers in the different subjects may only be partly explained by the limited availability of hardware and educational software in a particular grade or subject, but that other factors also contribute to the low use of computers in existing subjects.

The results of the Comped study may shed some light on the actual practice of staff development in the schools and its relation with computer use (or non-use) of existing subject teachers. In this paper we will describe the status of knowledge and skills (of teachers in existing subjects) in educational computer use, their involvement in teacher training, and we will relate these to the actual use of computers in the classroom.

Another aspect of staff development is the communication and interaction among teachers, about which some results will be presented as well. Although staff development can be studied on school and school transcending level (for example support to teacher training, the role of the principal in the innovation), this paper will be restricted to the teacher level.

Research questions and data source

The data to be used in this paper were collected in 1989 in the 'Computers in Education' study that was conducted under the auspices of the International Association for the Evaluation of Educational Achievement (IEA). Data were collected via questionnaires at three educational levels (elementary education, lower secondary education and upper secondary education) in representative samples of schools. Principals, computer coordinators and teachers of informatics, mathematics, science and mother tongue were included in the sample.

Twenty countries participated in stage 1 of this study (see Pelgrum and Plomp, 1991). In this paper, data will be presented across countries which means that individual differences within and between countries are not discussed.
The data collection will be repeated in 1992 (stage 2) and will -then- also be extended to measures of student outcomes. Details of the design (instrumentation and samples) as well as a description of results can be found in Pelgrum & Plomp (1991).
In this paper, we will focus on teachers of the existing subjects (in the Comped study: mathematics, science and mother tongue) in lower secondary education, leaving out teachers of computer education.
The following research questions will be addressed:

- What experience, knowledge and skills do teachers have concerning computers and their use in education, and how are these related to training?
- Is there a relation between training received and actual practice of computer use in the existing subject lessons?
- Which teachers participated in training activities and what topics were covered in their training?
- Is communication and interaction among teachers (being an aspect of staff development) an important indicator for computer integration?

The existing subject teachers who work with the computer can be divided into three groups. As measure for integration, Pelgrum and Schipper (1992) defined a 10-point scale indicating the number of topics in each of the existing subjects (maths, science and mother tongue) is which the computer is being used. For the purpose of this article, we define three categories: low integrative users (with 1, 2 or 3 points on this integration scale), the mediate integrative users (with 4, 5, 6, or 7 points) and the high integrative users, with 8, 9 or 10 points on the computer integration scale. These three groups and the group of non-using existing subject teachers will be compared on a number of variables (number of respondents: non-users 4684, low integrative users 887, mediate users 428, high users 212).
To answer the research questions, variables like the knowledge and skills of teachers, whether training was received or not, the amount of training received, content of training, the number of years a teacher works with the computer, the actual use of the computer in the class and issues related to communication and interaction were included in the analysis.

Results and conclusions

Years of instructional use

An analysis of the number of years teachers work with the computer for instructional purposes shows that more integration of the computer can be expected when a teacher works longer with the computer. Apparently, computer integration is such an complex innovation that it cannot be expected to be fully implemented within a short period of time.

Knowledge and skills

The results show that using teachers know significantly more than non-users and a significant difference on the self-rating scales can be reported for all pairs out of the three using groups, indicating that the group of high integrative users have

most knowledge and skills. Furthermore it is found that amount of knowledge and skills is related to training.

Training

The amount of training received and the type of topics covered in training are related to the amount of computer integration. A conclusion from these findings is that a close look at the types of teacher training is necessary in order to find out how much and which topics should be included. An indication about the importance of the content of training is gained in this paper, leading to the conclusion that mainly pedagogical / instructional aspects are very important to include in training. Because the results show that these aspects were only limited present in the training activities up till 1989, an implication from the results is that teacher training institutes should take these aspects into account.

Relation training and actual use

Teachers of existing subjects often teach also about computers. The computer topics they are covering in their lessons are correlated with the topics addressed in their teacher training (r=0.39).

Communication /interaction

Besides training, the influence of communication and interaction on computer integration was studied but no significant results on this relation was found. However, one may not conclude from these results that communication and interaction are not important in staff development. In the first stage of Comped, only a few variables concerning these aspects were included in the instruments and it might be that the operationalization of this aspects was not adequate. In stage two of Comped more detailed variables concerning this aspect are included in the instruments.

Concluding remark

This paper presented data across countries. It is useful to study national situations concerning staff development as well. Within the group of using teachers, it is interesting to trace in which country the high integrative teachers are located. Our data show that the group of high integrative users in our sample mainly come from Austria (10.8%), France (35.8%), New Zealand (9.9%) and USA (23.6%).

References

Fullan, M.G. & Stiegelbauer, S. (1991). The meaning of Educational change. New York: Teachers College Press.

Orlich, D.C. (1989). Staff development: enhancing Human Potential. Boston: Washington State University.

Pelgrum, W.J. & Plomp, Tj. (1991) The use of computers in education worldwide: results from the IEA 'Computers in Education' survey in 19 Education Systems. Oxford: Pergamon Press.

Walker, D.F. (1986). Computers in the curriculum. In: Culbertson, J.A. & Cunningham, L.L. (eds). Microcomputers and education, eighty-fifth yearbook of the National Society for the Study of Education. (pp22-40). Chicago: Chicago Press.

FACTORS DETERMINING THE DEGREE OF COMPUTER IMPLEMENTATION: A COMPARISON BETWEEN FIVE EDUCATIONAL SYSTEMS

Alfons ten Brummelhuis, & Albert Tuijnman, Centre for Applied Research (OCTO), Enschede, The Netherlands

Based on the theoretical framework of the Comped study, a hypothesized path model (LISREL) is developed in which important determinants of computer implementation are specified. The purpose of the analyses is to investigate the factors that may explain why some schools are using computers for educative purposes to a greater extent than others. Besides, the results show the interrelatedness of factors that influence computer implementation. The hypothesized model is tested with data derived from lower secondary schools in five educational systems: Japan, France, The Netherlands, Switzerland and the USA. This paper presents the hypothesized model and the over-identified recursive path model for each of the analysed educational systems.

EVALUATING THE IMPLEMENTATION OF ENGLISH, MATHEMATICS AND SCIENCE IN THE ENGLISH NATIONAL CURRICULUM: PROCEDURES AND EARLY FINDINGS

T. Russell, Department of Education, University of Liverpool, Liverpool, United Kingdom

Introduction

The National Curriculum was first introduced in England and Wales in September 1988 and is now operating for all pupils in state schools during the years of compulsory education(on (5 - 16). There will be nine subjects in all implemented as part of the National Curriculum. Three form the core as the subjects on which there is most time spent by pupils.

The three core subjects were introduced as part of the National Curriculum before other curriculum areas and it is these which are the focus of the evaluation studies. The National Curriculum Council commissioned three institutions to evaluate the implementation of English, Mathematics and Science. The three centres made their proposals independently. There are, therefore, interesting and important differences between the three approaches which reflect the different subject areas, the different contexts in which the subjects were introduced and the different research methods of the teams. This symposium is presented when the three projects are half way through their work. It will focus on the methodologies employed in each of the projects, only referring to interim results to illustrate points.

Summary

Three projects were commissioned independently by the National Curriculum Council of England (NCC). Contracts were given for the three research teams to evaluate the implementation of English, Mathematics and Science in the National Curriculum. The projects will each be half way through their two year programmes by June 1992.

The symposium will focus on the various approaches taken by the separate projects and the reasons why the different methodologies were adopted. Influences on the methodologies employed include; the issues identified by others as having a significant impact on the implementation of the National Curriculum, the traditions of teaching in the individual areas of the curriculum, the various stages of primary and secondary schooling, and, finally, the identification and interpretation of these issues by the research teams themselves.

It is hoped that the symposium will provide a valuable forum for the discussion of research methods in the evaluation of curricula in general.

EVALUATING THE IMPLEMENTATION OF ENGLISH IN THE NATIONAL CURRICULUM

A. Walsh, Cambridge Institute of Education, & A. Millett, Centre for Educational Studies, King's College, University of London; Members of the Mathematics Evaluation Team, United Kingdom

Introduction

The Evaluation project reported here is part of a rolling programme of monitoring which is being conducted by the national Curriculum Council. Specific issues concerning the teaching of English to 5-14 year old pupils have been identified by Her Majesty's Inspectorate of Schools, by NCC and by others.

Aim

The aim in undertaking this programme of evaluation is to ensure that the task the teachers are facing is fully understood. Questions to be addressed need to target whether or not the curriculum requires alteration, whether teachers' knowledge and understanding is appropriate and whether the attainment (assessment) framework is an accurate reflection of pupil progress.

Method

A variety of methods of data collection are being used to investigate these possibilities and the methodology of this evaluation project will provide the basis for this contribution to the symposium.

EVALUATING THE IMPLEMENTATION OF MATHEMATICS IN THE NATIONAL CURRICULUM

B. Raban, Department of Education, University of Warwick, United Kingdom

Introduction

Evidence from Her Majesty's Inspectorate reports and from preliminary National Curriculum Council monitoring studies has revealed areas of concern in the implementation of the National Curriculum for Mathematics, at ages 5 - 14. The aim of this research is to pursue a 'nested' set of investigations, each designed to address systematically the issues and questions which have been identified.
The nature of school and teacher planning and review procedures is one focus of attention, especially the extent to which teachers are basing these on National Curriculum documents, either those containing general advice, or the very detailed statements defining each of the different levels in the mathematics curriculum. The use made of published schemes of books and other materials as mediators of the National Curriculum will be investigated.
Problems in the implementation of selected topics are also being investigated. It seems likely that some of these are caused by lack of relevant subject knowledge,

especially in areas of new content. However, lack of resources in terms of teaching materials and equipment may contribute to this. The statements which define the curriculum may be inappropriate, either in their content, their clarity, or the level in which they are placed.
The extent to which teachers are meeting requirements in terms of process and content is being considered. Illustrative examples of promising practice in this regard will be another focus of the research.
Changes in practice since the introduction of the National Curriculum will be considered throughout the study.

Methods
A range of methods of data collection, including questionnaire, interview, teacher group development work and case study, is being used to investigate these concerns. The methodology of this evaluation project will form the central focus of the contribution to the symposium.

Results
The project outcomes are expected to be advice to the National Curriculum Council about possible modifications to the Mathematics Order and professional support which might he needed.

EVALUATING THE IMPLEMENTATION OF SCIENCE IN THE NATIONAL CURRICULUM

T. Russell, A. Qualter, & L. McGulgan, Science Evaluation Team, University of Liverpool, Liverpool, United Kingdom

Introduction
The Research rests on a number of issues which have been raised by HMI (Her Majesty's Inspectorate of Schools) by NCC (National Curriculum Council) and others. It has been observed that teachers are to some degree failing to cover certain aspects of the National Curriculum Science, they may be overlooking them, or they may be deferring them until later in the pupils' schooling. The reasons why this is the case are being explored with a focus on planning for teaching science. The second study involves a consideration of the match between the levels in the National Curriculum, intended to represent progression in learning, and the order in which pupils develop their understanding in Science. Finally we intend to consider the appropriateness of the National Curriculum in Science using a mixture of individual interviews of teachers and pupils, National Questionnaire, group interviews and classroom observation in meeting the not inconsiderable challenges of this project.
In attempting to deal with these issues it is necessary to explore the nature and extent of many primary teachers' confidence, background knowledge and experience in teaching science. It is also important to consider their approaches to planning and the extent to which their views of good primary practice coincide with their interpretation of how the National Curriculum should be implemented. The

extent to which INSET, colleagues, or the approach taken by the school influence the way a teacher implements National Curriculum also needs to be taken into account.

In the secondary school sector many of the same issues apply, however, others become important. The extent to which science departments have adopted balanced science in the lower school, may be an important factor in determining how the National Curriculum is being implemented. The approach to differentiation taken by the school and the department will be explored, including the approaches taken by individual teachers in planning and delivering their courses. It is also evident that the scheme of work adopted by departments both reflects their views of good science teaching and influences how they teach. Many schools have developed their own schemes and these along with published schemes will be included in the analysis.

Results

The project outcomes are expected to be of two kinds; Guidelines for the National Curriculum Council as to how to improve the National Curriculum; and Guidelines for the development of a coherent approach to the delivery of teacher training in implementing science in the National Curriculum.

CURRICULUM RESEARCH AND CURRICULUM PRODUCTIVITY: THE CONTRIBUTION OF CURRICULUM TO EDUCATIONAL PRODUCTIVITY

Wijnand Th.J.G. Hoeben, Institute for Educational Research (RION), University of Groningen, Groningen, The Netherlands

General summary

Curriculum research and educational productivity research are frequently seen as two separate research programs. Most curriculum research does not deal with productivity, but with issues of content, objectives and instructional strategies in the curriculum and with implementation of the curriculum. Educational productivity research deals with educational outcomes and the independent variables of instruction, of the school, and, sometimes, of different curricula that are used. However, to decide on curricular issues in order to provide or to induce effective instruction and effective learning to students, a conceptual link between curriculum and outcomes is necessary. Some form of causal reasoning from curriculum to educational outcomes may be even thought of as necessary in order to make productive curriculum decisions. We need therefor an theory of curriculum productivity.

The symposium discusses the possible causal linkages between curriculum and educational outcomes by reviewing curriculum research and productivity research in some school subjects, especially language education, mathematics education and science education. The reviews will be introduced by presenting a theoretical framework, that may clarify the issues at stake, and that will facilitate a systematic comparison of research in different subjects and different countries.

Reviews on curriculum research will identify issues of objectives, content, time, materials and instructional strategies. Different curricula may be interpreted as practical elaborations of different theoretical positions on these issues. Research on implementation is reviewed with special reference to instructional strategies and student achievements.

Reviews on productivity research will describe student achievements in the same school subjects and will review the available research data on causal explanations. Special attention will be given to instructional variables that explain achievements, and to practical configurations of instructional variables that may be thought of as operational curricula.

Discussants will challenge reviewers of curriculum research to make clear statements on productivity implications of the available research data. They will also challenge reviewers of productivity research to make clear statements on

curriculum implications of the available data. The discussion will focus on possible models of causal reasoning or theories that explain student achievements by curriculum factors that (may) have influence on achievement, and by factors of the implementation and instruction processes that (may) mediate the influence of curriculum factors.

A THEORETICAL FRAMEWORK FOR CURRICULUM PRODUCTIVITY

Wijnand Th.J.G. Hoeben, Institute for Educational Research (RION), University of Groningen, Groningen, The Netherlands

Curriculum, implementation, instruction and learning are the key concepts for building a theoretical framework that may explain curriculum productivity. Curriculum research has demonstrated some characteristics of the curriculum that influence its implementation, like explicitness, complexity, practicality, and relevance. Implementation research highlights the curriculum in operation. Educational productivity research combines variables of students' aptitude and students' environment with variables of the instruction in explaining student achievement. In a theoretical analysis the variables of curriculum implementation are compared with the variables of effective instruction, thereby demonstrating a conceptual overlap in the theoretical frameworks of both research traditions. This conceptual overlap is taken as a starting point for formulating a possible theory of curriculum productivity.
Like students' aptitude and their environment are conditions for instructional effectivity, factors of school effectivity are conditions for effective curriculum implementation. In cross-national comparisons the differences in educational structure and policies and in the external support of schools have also to be taken into account. These factors will be seen as conditions in a theory of curriculum productivity.

CURRICULUM AND IMPLEMENTATION IN LANGUAGE EDUCATION

Wolfgang Herrlitz, University of Utrecht, Utrecht, The Netherlands, & Jan Sturm, University of Nijmegen, Nijmegen, The Netherlands

The paper will review curriculum and implementation research in language education. It will identify issues of objectives, content, time, materials and instructional strategies. Trends in the discussions on these issues will be highlighted as differences in intended curricula.
Different curricula and textbooks may be interpreted as different practical elaborations of the positions that are taken in the discussions on the issues identified. Research on curriculum implementation links the trends in the

discussions and the intended curricula with the educational practice (curriculum in operation).
Research on implementation, therefor, is reviewed with special reference to instructional strategies and student achievements in order to build up a body of research that overlaps productivity research in explaining achievements. The discussion on the theoretical linkages between both research traditions will thus be stimulated.

STUDENT ACHIEVEMENTS IN SCIENCE EDUCATION

F. Michael Connelly, Ontario Institute for Studies in Education, Canada, & Joseph Menis, Bar-Ilan University, Ramat-Gan, Israel

These reviews on productivity research will describe student achievements in mathematics, science and language education. They will review the internationally available research data on causal explanations of achievements and productivity. Special attention will be given to 1. research of instructional variables that explain achievements and productivity, and to 2. research of curriculumimplementation, that is in a terminology that foreshadows a viable theory of curriculum productivity, research of practical configurations of instructional variables that may be seen as more, respectively as less effective operational curricula.
As far as is possible, the results will be interpreted in terms of differences between curricula on the respective subjects in order to build up a body of research that overlaps - as much as possible - with research of curriculum and implementation. So doing the reviews will contribute to the discussion of the possible causal interpretations of curriculum research in terms of educational productivity research, and, ultimately, to a discussion of the possible contributions of curricum to educational productivity.

Discussants

Jan van den Akker, University of Twente. The Netherlands and Jan Terwel, University of Amsterdam, The Netherlands

The discussants will challenge reviewers of curriculum research to make clear statements on productivity implications of the available research data. They will also challenge reviewers of productivity research to make clear statements on curriculum implications of the available data. The discussion will focus on possible models of causal reasoning or theories that explain student achievements by curriculum factors that (may) have influence on achievement, and by factors of the implementation and instruction processes that (may) mediate the influence of curriculum factors.

STRUCTURE OF EDUCATIONAL RESEARCH

T]. Plomp, Department of Education, University of Twente, Enschede, The Netherlands

BERA AND EDUCATIONAL RESEARCH IN THE UNITED KINGDOM

Michael Bassey, & Janet Powney, British Educational Research Association (BERA), United Kingdom

The organisation of educational research in the UK

The scale of educational research in the UK is difficult to judge. There is a relatively small number (perhaps 300) of full-time researchers in research institutes (NFER, SCRE), universities, polytechnics, and some local education authorities; a large number of teachers in higher education engaged in teacher training who do some school-based research; and a large number of other teachers in higher education in non-education disciplines who do research into their own teaching. According to the NFER biennial index there are 2300 of the above listed people engaged in educational research in the UK. Beyond this is an unknown number of school teachers who engage in action research in their own schools - at any one time there must be at least 1000 of these studying for master or doctorate awards of some of the 60 universities and polytechnics which have departments, schools or faculties of education.

There are two major research institutes: the National Foundation for Educational Research [NFER] (in Slough, south of London) (about 40 projects at a time), and the Scottish Council for Educational Research [SCRE] (in Edinburgh) (about 30 projects at a time).

There are a number of associations for educational researchers, including:

- British Educational Research Association [BERA] - all forms of educational research, learned society, national membership;
- Scottish Educational Research Association [SERA] - all forms of educational research, Scottish membership;
- Educational Research Network of Northern Ireland [ERNNI] - mainly practitioner research, Northern Ireland membership;
- Classroom Action Research Network [CARN] - action research, national membership;
- London Education Research Network [LERN] - mainly quantitative research linked to local authorities in London;

- South West Action Research Network [SWARN] - action research in the South West of England;
- Society for Research into Higher Education [SRHE] - policy and practitioner research into higher education, national membership;
- In addition a number of curriculum associations and specialist societies hold research conferences from time to time.

The British Educational Research Association

The British Educational Research Association (BERA) was formed in 1974 with "the broad aim of encouraging the pursuit of educational research and its applications for the improvement of educational practice and the general benefit of the community." The Association has about 600 members who pay an annual subscription of £35.

The major activities of the Association are as follows:

Publications

- British Educational Research Journal Editor: Dr Gaby Weiner. Produced by Carfax Ltd. About eight refereed papers per issue and book reviews. The 4 issues of 1992 will form volume 18;
- BERA Dialogues Editor: Donald McIntyre (previously Dr Pam Lomax). Produced by Multilingual Matters Ltd. Occasional publication. Refereed papers around a theme, eg Number 6 (1992) *Local Management of Schools: Research and Experience*, edited by Dr Gwen Wallace;
- BERA Research Intelligence Editor: Dr Pam Lomax. Quarterly newsletter giving substantial reports on BERA activities, announcements of educational research meetings throughout the UK and of some international events, and some articles and book reviews.

Meetings

- Annual Conference. A four day residential conference at different centres throughout the UK held at the end of the summer. It is the largest annual gathering of educational researchers in the UK. Last year at Nottingham there were 21 symposia (each including about 5 papers - topics included Action research, Local management of schools policy, School governors, National curriculum policy, Assessment, School improvement, Teacher education, Supply teaching, Transition from school to work, Library use, Interface between qualitative and quantitative research) and 40 individual papers. This year it is at Stirling, in Scotland.
- Regional Seminars. Over the last year there have been 10 of these one day events, held around the country at Edinburgh (National curriculum assessment), Sunderland (Change in classroom practice), Bristol (National curriculum assessment), Nottingham (Race and gender), Bath (Action research), Kingston upon Thames (Action research), Loughborough (Teacher's professional knowledge), Bournemouth (Action research in continuing education), Newcastle (City technology colleges) and London (Local management of schools).

Policy task groups

BERA Policy Task Groups are established by the Executive Council to organise the discussion, review, and critique of policy research; to encourage and promote

the dissemination of policy research; and to endeavour to provide opportunities for policy makers to be appraised of research findings pertinent to their decisions. At present there are 5 policy task groups: National curriculum, Assessment, Local management of schools, Teacher education and Adult and continuing education. Various publications and meetings have emanated from these groups.

Representing educational research nationally

As a learned society BERA is invited from time to time to make submissions to government and national bodies on research issues. During the last year BERA has made submissions to the Economic and Social Research Council (ESRC), the Universities Funding Council and the National Commission on Education. BERA is active in the UK Association of Learned Societies in the Social Sciences [ALSISS]

BERA's synoptic view of educational research

BERA sees educational research as attempts to advance educational knowledge through systematic enquiry and critical debate and governed by three ethical principles: respect for truth, respect for persons, and respect for democratic values. Respect for truth means that falsehood is unacceptable, respect for persons means that people should not be abused, and respect for democratic values means that researchers should be free to enquire and to publish their findings, subject to the other two ethics.

BERA takes a broad view of educational research to include any disciplined enquiry - empirical, reflective and creative - which serves educational judgements and decisions.

This may be conducted in educational settings such as nursery, primary, secondary, further, higher, continuing and adult education; industrial, commercial and professional training; and local and national systems of education. The methods and techniques of inquiry may originate from a practical view of how knowledge is best generated and utilised by teachers, and educational policy makers and managers, or they may draw on the methodologies of other social science disciplines, such as sociology, psychology, anthropology, philosophy or economics.

Empirical research means the kind of research where data collection is centre stage; where data is systematically collected by strict procedures, critically analysed, interpreted and conclusions drawn. *Reflective research* refers to systematic and critical thinking in which the findings of empirical research are the starting point for review and argument about educational issues. *Creative research* means the devising of new systems, the development of novel solutions, and the formulation of new ideas, by systematic and critical enquiry. In practice, of course, these three realms often overlap.

BERA's members embrace many forms of educational research. They include:

- research workers who seek to understand educational situations, those who set out to evaluate them, and those who strive to use research methods to implement change in them;

The first group seek to describe, interpret or explain what is happening without evaluating or inducing any change. They are trying to portray the topic of the enquiry as it is, without disturbing it, and to give a theoretical account which links

with existing theoretical ideas. Usually they are outsiders, ie people not engaged directly in the educational practice under study.
The second group evaluate what is happening in relation to their own educational values or those of the sponsors of the enquiry. Usually they are outsiders engaged in a collaborative venture with insiders. They are trying to portray the topic of the enquiry so that the insiders have a knowledge base from which to implement change.
The third group use research approaches directly for implementing change. They are insiders engaged in classroom teaching or in administration who intervene in the ongoing action (hence the description "action research"), monitor and reflect on the consequences, and introduce further change on the basis of their systematic and critical study.

- research workers who search for generalisations and those who study singularities;
  The search for generalisation requires the investigation of large populations, usually studied by appropriate sampling, and by intention leads to statements which can be used to predict what will occur in other situations.
  The study of a singularity is investigation of something within a local boundary of space and time; it cannot be used to predict probabilities, but if it can be related to other situations, it may be valuable in suggesting possibilities for future action in those other situations.
- research workers who have different beliefs about the nature of reality, for example the positivist paradigm and the interpretive paradigm.
  Researchers working within the *positivist paradigm* see reality as separate from themselves and expect investigators to have the same perceptions of shared phenomena and thus common understandings.
  Researchers working within the *interpretive paradigm* see reality as a social construct and so do not necessarily expect other investigators to have the same perceptions or understandings of shared phenomena.

BERA's members hold different beliefs about the nature of education itself. Taking, for example, the statement that *education is both the nurture of personal growth towards worthwhile living, and the conservation, transmission, development and renewal of worthwhile culture*, there are widely different beliefs about what constitutes "worthwhile living" and "worthwhile culture", about the relative importance of these two, and, indeed, about the merit of the statement itself.

Five issues of concern in the UK

- Much educational research in the UK aims to inform the practice of teachers and the policy of politicians, managers and administrators, but we find great problems in communicating findings effectively to those who might use them. This is presently exacerbated by a political climate in which professional expertise is often scorned;
- Much educational research in the UK is carried out by university and polytechnic teachers in teacher training and by local education authorities The policy of the present Government is to transfer much teacher training to schools and to drastically reduce the local education authorities. Hence in the next few years

there may be a substantial reduction in the amount of educational research being carried out.;
- For full-time researchers (he contracts are short term, pay is low, and the career prospects are such as not to encourage this work;
- The expansion of higher education student numbers is reducing the time for staff to engage in research, but university funding of research is now competitively based on the extent and excellence of research. Hence research may cease in some institutions;
- There is growing evidence that some of the Government agencies which support educational research are reluctant to allow researchers to publish their findings. Secrecy is counter to the ethic of democratic values.

EDUCATIONAL RESEARCH IN THE NETHERLANDS

Prof. dr. Gellof Kanselaar, President of the Dutch Educational Research Association (VOR), Utrecht, The Netherlands

Introduction

Firstly I'll give some information on the Dutch Educational Research Association and on the education support structure in the Netherlands.
Secondly I'll describe the way in which research is organized in the Netherlands.

The Dutch Educational Research Association (VOR)

The Dutch Educational Research Association (VOR) was funded in 1975. About 450 persons are a member of the VOR. The VOR consists of seven divisions:
- Division Curriculum Studies and Organization;
- Division Learning and Instruction;
- Division Education and Society;
- Division Methodology and Evaluation;
- Division Teacher Training and Teacher Behaviour;
- Division Human Resource Development: vocational training and in-company training;
- Division Higher Education and Adult Education.

Activities of the VOR are among other things:
- The responsibility for the organization of a yearly conference on educational research. About 600 people attend this conference which is organized at a different university each year.
- Publication of the Journal of Educational Research (TOR) bimonthly

The Education Support Structure

Agencies for education research, development, dissemination and guidance in the Netherlands together constitute the national support structure for education. It was recognized at an early stage that the different support institutions would have to function within a coherent structure in order to offer the most effective support to the development of schools and schooling. This need for a coherent structure became a phenomenon in its own right and was one of the principles on which

legislation concerning the support structure was based. The Education Support Structure Act was the result of discussions over many years between school organizations and the government about how coherence in the activities of the support agencies could best be realized.
The education support structure in the Netherlands is made up of two types of institutes: those that provide general educational support and those that have a specific mission. Among the general educational support institutes are some 65 school counselling services (SBDs), which operate at a local or regional level. Services offered by the SBDs range from individual psychological testing and advice to helping schools develop remedial teaching programs and local instructional materials. In general SBDs are of a "neutral" character in the sense that they are not affiliated to, or represent any of the religious denominations (Catholic, Protestant, non-denominational) that play an important role in many aspects of Dutch education.
In addition, there are three national educational advisory centres (LPCs). Unlike the SBDs, these offer their services (throughout the country) only to schools of the particular denominations they represent. The services offered by these centres focus on secondary education and on structural and organizational problems.
The second type of educational support system has a more specific mission. It consists of three national institutes which focus on the development and improvement of the education system in research, measurement of educational achievement, and curriculum development. These institutes are respectively, the National Institute for Educational Research (SVO), the National Institute for Educational Measurement (CITO) and the National Institute for Curriculum Development (SLO).
CITO develops and administers tests for national examinations and for individual schools. It processes test results and reports the outcomes to schools, parents and pupils. For two years CITO, together with SVO, has also been engaged in the national assessment of educational achievement. CITO is responsible for developing and administering the assessment instruments and for reporting the results, while SVO conducts the research on the basis of which the findings of the assessment programs are to be elucidated.
SLO's task is the development of general curricula for the schools. The adjective "general" should be stressed here, because, due to the constitutionally anchored principle of freedom of education in the Netherlands, a national and publicly financed institute like SLO cannot prescribe in detail the form and content of school curricula. Therefore, SLO develops a broad curricular framework for almost every subject area in primary and secondary education in such a way that individual schools can implement the curricula in accordance with their own educational philosophy.

Organization of Research in the Netherlands
The research carried out in the Netherlands falls into four categories, according to the sources through which it is financed.

Research financed by universities
The primary financial source is called the 'Conditionally Funded Research Program' (VFO). It covers the right to research time granted to university lecturers

who are on the permanent staff and participate in a research program. About 40% of their time is available for carrying out research if they are a member of an officially approved research program. These research programs are evaluated every five years. Criteria for evaluation are output (publications) and coherence of the program. There are about 15 conditionally funded research programs in educational sciences in the Netherlands.

Research funded by the Institute for Educational Research (SVO)

The Institute for Educational Research was instituted in 1965 by the Dutch Ministry of Education and Science as The Foundation for Educational Research (SVO, after its Dutch initials). As a result of the Education Support Act (WOV) in January 1987, SVO became part of the Education Support Structure under its present name.

- The Education Support Structure and SVO

SVO can be compared to a national Science Council on the one hand, and to a research contracting agency on the other. The leading principle in SVO's activities is that of the customer-contractor relationship. The main customers are the nations' four Schools Councils, on which both the school practice and the national government are represented. Both parties are entitled to 40 % of SVO's annual budget of approximately 7 million ECUs. The remaining 20 % is used for "fundamental strategic research" on SVO's own initiative. This latter type of research is oriented toward theory development and evaluation.
Most policy and practice oriented research is carried out by five university-based educational research institutes. About 75 % of all educational research in the Netherlands is channelled via SVO. However, it is part of SVO's task to co-ordinate all educational research. Decision-making takes place through SVO's Board of Directors, on which educational practice, the government and the research community are all represented. It is SVO's explicit responsibility to establish research programs for policy and practice oriented research.

- Documentary Support

Apart from programming, coordinating, assessing and funding research, SVO offers documentary assistance. SVO coordinates all documentation activities for educational research within the national educational research within the national education support structure. The Documentation Department enters research findings in the national Dutch database DION. SVO also participates at several levels in the European Documentation and Information (EUDISED) of the Council of Europe.

The Dutch Organization for Scientific Research (NWO)

NWO is the organization for fundamental research. NWO consists of 34 foundations, one of which is the foundation for pedagogical research. One of the five divisions of this foundation concerns research in education. About 5% of the budget is spent on research in the social sciences (9 million ECU) and only _ 100.000 ECU is spent on educational research.
Individual researchers may submit a research proposal on their own initiative through a university department or a research institute. This will normally concern research that is of a fundamental character. The proposal is judged by peers.

Miscellaneous organizations
The fourth source constitutes a variety of funding agents, including government grants from the Ministry of Eduction and Science and other Ministries, local authorities and industry. In the last few years research groups execute an increasing number of research-projects from funding agents in the educational area itself on account of the decentralisation of governmental policy.

Discussion
One of the main problems for educational researchers is that it is very hard to get money for fundamental research, because researchers cannot submit a research proposal on their own initiative to the SVO. Most of the money is spent on applied research. The consequences are less theory building and short-term solutions for educational problems.
A second problem is the decrease in money spent on educational research by the Ministry of Education and Science.

233

CHAPTER 3

LEARNING AND INSTRUCTION

LEARNING AND INSTRUCTION

Sanne Dijkstra, & Robert-Jan Simons

The nearly 100 contributions to the theme Learning and Instruction are categorized in different subthemes. The coordinators decided to first make a main categorization of papers in those that contribute to general theories of learning and instruction and to the description of tools for instructional design and those that address learning and instruction within domains. The first main category encompasses contributions to motivation and learning, metacognition, learning style and cooperative learning, whereas the second category comprises papers on science and mathematics education, reading and writing instruction, foreign language teaching, social skill training, computer programming instruction and teaching computer applications.

The categorization sometimes may seem arbitrary. Those who prepared papers about their educational work in domains such as mathematics and science use theories from educational psychology as a framework. General theories of learning and instruction and their empirical validation are presented and discussed in 6 symposia. The strong ties between psychology and education are highlighted in symposia on motivation and self-regulation (Boekaerts and Simons), on the relation between cognitive and metacognitive skills (Veenman and Samarapungavan) on teaching inductive thinking skills to disabled and normal students (Bächel). Further one symposium addresses the role of domain-specific knowledge and expertise in learning and instruction (De Corte and Dochy). The influence of prior knowledge on learning is reviewed and discussed in a symposium prepared by Kirschner and Dochy. Recent developments in instructional design are presented and discussed in a symposium which is organized by Dijkstra. Besides the symposia several papers and posters on motivation and self-regulation and especially on metacognition will be of interest for those attending the conference. The contributions on metacognition address the relationships between metacognition and conceptual growth as well as improvement in skills. Several papers present the results of research on instructive environments for the acquisition and practicing of reading and writing skills. Some studies emphasize the fundamental processes, such as word recognition, whereas other studies present the effects of instructional variables on learning to read and write. A substantial amount of papers deal with mathematics and science teaching. Steiner's keynote address is devoted to the psychology of mathematics education, especially semantic network theory applied to construction of algebraic-mathematical micro-networks. His presentation will be of strong interest both for those who work in educational psychology settings and for those who work in the

domain areas. The papers on mathematics and science education further address issues which are on research agenda's either for elementary, secondary or university education. Some papers, however underline the continuity between learning in one institution and the next. Salient are those papers which address the development of science concepts of elementary school students, such as energy and conservation of matter. In mathematics education attention is paid to realistic mathematics instruction, a strong trend in instructional development in this domain. The papers on the teaching of computer science mainly address the meaning of programming for the acquisition and practicing of problem-solving strategies and skills. The use of logo is still popular.
The coordinators noticed a general trend towards studies of instructional strategies within several subject matter domains.

TEACHING ENERGY AND PRIMARY SCHOOL CHILDREN'S ALTERNATIVEFRAMEWORKS

Ricardo Trumper, Oranim, School of Education of the Kibbutz Movement - University Division, Israel

Introduction

When should we start teaching energy? Warren (1986) claims that the energy concept should be taught only to students who have reached a high level of abstract reasoning. As opposed to that, Solomon (1986) claims we should start teaching energy in primary school. Children's alternative concepts about energy appear to arise from a context which is inappropriate for school science. A student can read a statement in science, and assimilate it by using the everyday interpretation of the word. This is not surprising since children are faced, through the media, with assumptions about how things move and about their energy, which are at odds with scientific concepts. So, when should we start teaching energy? At which age do children construct the relevant frameworks, which can be used ,as building blocks in the teaching of the concept? (Trumper 1991). A longitudinal study dealing with this issue will be described in the next section.

Research method

The study encompassed 398 students, aged 10 to 14, drawn from two rural regional schools in Israel. Students' energy preconceptions were analyzed by means of a questionnaire
including 3 tasks. Students were asked to:

- Write their first three associations with energy.
- Choose one definition of energy from five.
- Choose 3 of 8 pictures involving the energy concept and explain their choice using the word energy.

Children's definitions and explanations about their choice of pictures were classified according to the alternative frameworks defined by Trumper (1990) and Watts (1983).

Results

Analysis of associations

Associations held by younger students led to statistically significant differences. The main differences were:

- Elementary school students (grades 5 and 6) held many anthropocentric associations.

- The younger children (grades 5 to 7) had a broader range of associations. A chi square analysis showed no significant differences among associations held by students in grades 7 to 9.

Analysis of pictures chosen

Chi square coefficients among pictures chosen by children in the five grades were calculated and major differences were noted. There is a great difference between the pictures chosen by fifth graders and those chosen by all other students. Fifth graders chose pictures with human beings more often than their older schoolmates. The analysis of data from grades 6 to 9 showed no significant difference among them.

Analysis of alternative frameworks

Children use mainly four alternative frameworks ("cause", "product", "anthropocentric" and "deposit") to describe the pictures they have chosen. Chi square coefficients among these dominant frameworks were calculated and results showed significant differences among grades. Again, there is a major difference between the alternative frameworks used by fifth graders and those used by all the other students participating in this study. Young children are much more anthropocentric than their older schoolmates. The analysis of data from grades 6 to 9 showed no significant differences among them.

Analysis of definitions

Definitions were included only in the questionnaire presented to junior high school students. Chi square coefficients among definitions chosen by students were calculated and major differences among them appeared. The main difference is between definitions chosen by ninth graders and those chosen by students in the other two grades. Ninth graders preferred the "product" framework definition while seventh and eighth graders preferred the "cause" framework definition. No significant difference emerged from the analysis of data from grades 7 and 8.

Discussion

In Israel the energy concept is taught in 9th grade. The curriculum approach was clearly stated by Shadmi (1984):

There is some typical energy for each phenomenon; whenever there is an interaction between two phenomena, the process can be described by means of energy transformations - during all these transformations, the quantity named energy is being conserved.

The findings of this study support a teaching strategy in which students' understanding of energy should be extended from its human-centered beginnings to a more general notion gradually approaching the concepts of energy transformations and energy conservation. If a spiral curriculum is adopted, this strategy should be presented from 5th grade on, according to the following stages:

- Young children should be led from their anthropocentric framework to a more generalized framework wherein human beings are seen as energy agents in processes of energy transformations.
- Students should analyze a great amount of physical processes and describe them in terms of "causes" and "results".

- Students should get acquainted with as many as possible different forms of energy.
- The concepts of energy transformations and energy conservation should be introduced, on the basis of Shadmi's (1984) guiding principles cited above.

In such a way, the scientific domain of knowledge concerning the energy concept may be reinforced step by step, and lead eventually to the necessary abstractions needed at higher levels.

References

Shadmi, Y. (1984). An Outline of a Mechanics Course Based on the Israeli Junior High School Physics Curriculum. In A.Mayer and P. Tamir (eds.), Science Teaching in Israel - Origins, Development and Achievements. Hebrew University, Jerusalem, Israel.

Solomon, J. (1986). When should we start teaching physics? Physics Education, 21, 152-154.

Trumper, R. (1990). Being constructive: an alternative approach to the teaching of the energy concept - Part one. International Journal of Science Education, 12, 343-354.

Trumper, R. (1991). Being constructive: an alternativeapproach to the teaching of the energy concept - Part two.International Journal of Science Education, 13, 1-10.

Warren, J. (1986). At what stage should energy be taught?Physics Education, 21, 154-156.

Watts, D. (1983). Some Alternative Views of Energy. Physics Education, 18, 213-217.

TWO LEVELS OF SEQUENCING IN GERM LEARNING: COURSE AND TASK LEVELS

H. Vos, Department of Electrical Engineering, University of Twente, Enschede, The Netherlands.

Abstract
A university course of practicals has been redesigned with the aid of fundamental principles we call germ learning. Models for knowledge construction guide the sequence of learning. The objective of the course is to learn a generality concerning the way new knowledge (including skills) is constructed and validated. The generality is introduced in a domain of prior knowledge, practised in a new domain and applied in another domain (transfer). In each domain new skills (measurements, calculations) are constructed, elements of which are provided. The results and the learning processes of good, moderate and weak students have been evaluated. It turns out that weak students fail in the last domain in which transfer is required.

Introduction to the problem
The models we use for knowledge construction in germ learning include a concept model, a development model and a validation model. From these models a sequence of learning has been derived on two levels. On a global level (the course as a whole) the sequence of learning to acquire and validate knowledge can be designated to consist of three parts: introduction (orientation), development (practise, integration), application (transfer). On a session level the students have moreover to construct new measurements and calculations. On this level we distinguish: homework (theory, lab-guides and assessment questions), preparatory tasks, (confronting) measurement tasks, open ended problems. Our question was what kind of influence this sequencing has on results and learning processes.

Theoretical framework
Germ learning (Vos, 1991) combines the ideas of Ausubel about a true advance organizer with those of Davydov (1968) about a genetic germ. A germ of learning in our sense is a cognitive structure, part of the student's mind, that has to be found and developed by the student. Teaching with the aid of models helps to accentuate this cognitive structure. The level of thinking is raised by special tasks to reconcile contradictions inherent in the models. Specification of different problem situations helps to differentiate problem solving actions for constructing

the missing conceptual knowledge. Integration of actions is aimed at by the requirement of validated knowledge and by open (ill structured) problem solving tasks.

Method

The experiment consisted of a redesign of a university course of practicals implemented in 1986/87. This course consists of nine sessions of 3.5 hours, partitioned in three parts of three sessions each: introductory measurements, time domain, frequency domain. In the introductory part students learn how to measure, how to calculate and how to validate knowledge in a domain of prior knowledge. In the time domain the students practise what they have learned. Finally in the frequency domain the students apply what they have learned to a new domain. The content and objectives of the course remained unchanged. In the first session, in which an entrance test and remedial exercises were introduced, prior knowledge is needed. This prior knowledge is summarized in a short description and pictural representation of the functions of the apparatus. A reading task and a confronting task in handling the measuring apparatus have to be fulfilled.

The tasks of sessions two to six have been analyzed and more clearly separated according to function. Homework tasks intend to teach the students to acquire and assess information they need. This information involves knowledge about concepts and skills, as well as knowledge about the validation of knowledge. We inform the students that measurements and calculations have to "fit" and we provide them with a "method of inquiry" containing about 50 action elements grouped into the headings: problem analysis, acquiring information, selecting a hypothesis, experimental validation, conclusions, reporting.

Preparatory tasks teach the student to construct new skills (constructions, measurements, calculations) for the determination of a quantity. Measuring tasks intend to compare measurements and calculations of a given quantity, and to validate the outcome. These tasks are the main objective of the course. Open investigation tasks leave the students the choice for the quantity to be determined, as well as for the method to be used. These tasks enable the students to practise above the final level required.

The tasks in the last part (sessions seven to nine) remained the same in order to monitor transfer.

During the sessions the students learn to make logbooks: they have to write down what they think, what they do (calculations and measurements, handling apparatus) and what comes out of it (their results). These logbooks are assessed by the teaching-assistants after every third session of the course. The logbook of each part is marked. Final marks are mainly based on the last two logbooks.

In 1990/91 an intensive evaluation of our course was carried out, including a comparison with earlier data. The years before 1987 provided quantitative data with respect to the marks of the students, and qualitative data with respect to the instruction and learning process. In order to get quantitative data concerning the learning process questionnaires were used after each session in 1990/91 to evaluate homework and laboratory activities of the students. The teaching-assistants used scoring-lists along with their own assessment lists. The scoring can be considered a content-analysis of the written work of the students. We did observations to check our data.

Our study can be considered as a time-series-design. The situation before and after the implementation were compared. The students were divided into three groups (good, moderate, weak) and the data for these groups analyzed. In order to check how these students behave in other courses, we analyzed correlations and factors in the marks for the first years courses.

Results

Overall failure rates dropped from 52 % to 21 % (average of the four years before 1987 and after 1987 respectively). Marks are rather independent of the assistant grading the logbooks. Marks correlate better with marks in other courses than before. The time used for homework did not increase. Student interaction is large, both in the old and in the new situation. Student interaction is effective in most groups in the new situation. About half of the students fails just as previously in one group with rather few good students. Students get remarkably less assistance in the new situation.

Scores rise from the first to the second logbook, and are generally lower for the third one. No complaints about the handling of apparatus have been found in the new situation. Students find it difficult to use the "method of inquiry". In the unchanged last part of the course some students are still complaining as before that they have to apply new not yet fully treated concepts in the laboratory, but most of them now are able to fulfil their tasks.

The marks for the logbooks show that good students perform usually best. With respect to measuring tasks they show a rise in performance in the second part and a decrease in the third one. Moderate students get nearly level with the good ones, starting at a lower level. They also show a decrease in the last part with respect to measuring skills, but are steadily improving their performance on study skills. The weak students fail in the last part, although their second logbook scored higher than the first one. Good and moderate students show better performance in selecting the difficult, relevant homework questions and they use much more time for these questions. Good students ask more often short help of the teaching assistants, weak students least.

Observations gave supplementary results only.

Conclusions

Prior knowledge needed has been activated satisfactorily in the first session, and has been upgraded. This is considered to be development and integration of a germ after partial knowledge has been acquired.

The results of the students improved satisfactorily. In groups with strong interaction it is probably of importance that the group contains not too few good students. Tasks showed more variation and included more difficult ones. Therefore mastery learning cannot be the explanation of our results.

Most students have learned to acquire by themselves the information needed in the unchanged last part of the course, in which transfer to a new domain is needed. They have also learned to compare results of measurements with results of calculations. Therefore it is concluded that on the global level, most students have learned how to acquire and validate knowledge (to apply the required knowledge construction).

Weak students fail in the last part of the course. Weak students have learned something in the first and second part. The conclusion is that our course selects on transfer. This is considered an indication that a germ has been accentuated and developed by most students.
Most students pass the course. Thus students learn more than before, more students learn it and in less time. Our education has shifted from remedial teaching (reparation of lacking skills afterwards) to preventive teaching (giving students the opportunity to prepare themselves effectively and efficient).

Discussion

The teaching-assistants give during the sessions minimal help (in the new situation). Superficially viewed we would say, as students and staff do, that assistants should take more initiative for help. A better conclusion is that the behaviour of the assistants contributes to learning to acquire information. This seems to be a confrontation in instructional design itself. The conclusion is probably valid only when the information provided is complete and interaction in the group is strong. In groups with too few students it is suggested to strengthen the group with some good students after a short entrance test.
Observations turn out not to be a good instrument for our purposes. Turning knobs (each apparatus has about 20 rotational knobs or switches) probably requires "turning-aloud-protocols" in order to study this cognitive-motoric skills.
Observations gave qualitative data with respect to the interaction of the students as a group.
Effect of the models has not been proved directly. The models help to design confronting tasks and tasks of a greater variety, to sequence the learning process on two levels, and, last but not least, to better understand the educational implications of the objectives of their own course.

References

Davydov, V.V. (1968). The relation between abstract and concrete knowledge in education (in Russian). Voprosy Psychologii 6, 34-50.
Vos, H. (1991). Learning and transfer: the use of germ learning (in Dutch). Tijdschrift voor Onderwijsresearch, 16, 261-278.

PROBLEM SOLVING IN SECONDARY PHYSICS AND SUCCESS IN HIGHER EDUCATION

Monica Fergusson-Hessler, & Ruurd Taconis, Physics Education Group, Eindhoven University of Technology, Eindhoven, The Netherlands

Abstract

This study investigates the problem solving skills required in first year physics courses, but not being attended to in secondary physics. Explicit and specific problem solving skills were elicited from 'stumbling-block' examination problems, and compared to skills required at secondary level. Test problems were constructed, requiring first year skills, but only secondary level physics. First year students were tested on these problems, and individual scores were given for mastery of specific skills. These scores were related to examination results, and significant correlation coefficients were found both for grades in individual subjects, and for total examination results. Possibilities will be investigated of including some of these skills in teaching material for the highest forms of secondary education.

Introduction

Difficulties experienced by first year students in higher vocational and university education seem to be caused by lack of problem solving skills rather than by lack of subject matter knowledge (van Dijk, 1990). This study investigates the problem solving skills required in higher education, but not being attended to in secondary physics education, and discusses the possibility of including attention for these skills in teaching material for the highest forms of secondary education.

Extensive research on problem solving has led to the conclusion that the cognitive skills involved in this activity are closely linked to the knowledge base of the problem solver (Schoenfeld, 1985; Alexander & Judy, 1988; Perkins & Salomon, 1989; de Jong & Ferguson-Hessler, 1991). A problem solving strategy consists of a number of stages or cognitive activities such as analysis (structuring information and adding new elements of information from memory), planning (selecting laws, concepts, and procedures from memory), and checking (comparing the result with expectation and known facts). Also, while the main steps of the strategy, like analysis and planning, are general skills, other (sub)steps are more or less strongly bound to the subject matter, like the supporting process 'defining the system' in Mechanics or Thermodynamics.

244

Among the questions investigated in this study are (Taconis, Stevens, & Ferguson-Hessler, 1992):

- Which problem solving skills are required for the solution of examination problems in Physics at the moment of leaving secondary education?
- Which problem solving skills are required for typical examination tasks in the first year of higher education?
- Is mastery of this type of skills at the moment of entering higher education related to success in the first year?

Method

The study was carried out in two steps: identification and description of problem solving skills required in secondary and higher education, and search for relations between mastery of these skills and measures of success in higher education. Subject matter domains were Mechanics and Electricity and Magnetism.

From a number of institutions teaching physics at tertiary level, first year examination problems were collected, which had turned out to form 'stumbling-blocks' for the students, that is to say that less than half of them had been able to solve the problems. For each of these problems an 'Ideal Solution' (IS) was worked out, representing the optimal solution the teacher expected. Using a five-stage strategy for problem solving in physics, developed in earlier research (de Jong & Ferguson-Hessler, 1984), we analyzed the IS's of these problems in terms of specific cognitive problem solving skills, and collected a list of skills which were necessary and/or useful for the solution. The same type of analysis was carried out for a number of leaving examinations for two types of secondary schools: HAVO, which prepares for higher vocational education, and VWO which prepares for university.

By comparing two pairs of lists (HAVO with higher vocational education and VWO with university education) we identified the skills which were not required at secondary level but did play a central role at first year level. Some of these skills were strongly dependent on subject matter knowledge which is typical of higher education, for instance selecting a contour or a surface for the application of one of the maxwell equations. A number of skills, however, were of the general type, and not bound to subject matter, for instance 'collecting information' or 'analysis of the given problem situation', and could well be incorporated in problems covering secondary level physics only. A number of specimen problems of this type were constructed, and validated as to level and type of skill by experts.

In order to investigate the relation between mastery of the cognitive skills identified and success in examinations, we constructed two tests, each consisting of a selection of specimen problems, and gave each of the tests to approximately 100 first year students in university and vocational higher education respectively, at the beginning of the academic year. Students were asked to solve the problems in writing, and to give information on their arguments and choices. Their solutions were analyzed and compared to the IS, and scores were given for the presence and correctness of the problem solving skills needed. A separate score was given for each of the following categories of problem solving skills: collecting information, analysis, planning, supporting processes, and calculation.

These scores were considered to measure the mastery of relevant skills at the beginning of the first year of higher education. Relations between these scores on

one hand, and leaving examination results for physics, and first year examination results on the other hand were investigated. Several measures were used for success in first year examinations: grade per subject (GRADE), percentage of passes in first examination period (PASS), and quartile of average result of this period (SCORE).

Results

A few general skills were found to be required at both levels investigated, and no sharp distinction can be made between general problem solving skills needed in higher education only and skills needed at secondary level as well. There are, however, clear differences in the number and quality of the skills required: in higher education these are more frequent, more complex, less well known to the students, and have to be carried out with less support than at secondary level. Most problem-type specific skills were found in higher education only. One exception was the supporting skill of selecting a set of coordinate axes to be used for the solution, a skill common to a number of types of problems.

In relating the variables measured, we confirmed a correlation, which is well known for physics in the Dutch educational system, that between leaving examination results and success in the first year of higher education. Pearson correlation coefficients found for different measures of success range from .46 (GRADE) to .63 (SCORE), with $p<.01$ in all cases. Individual skills scored showed correlation coefficients with different measures of success ranging from low and not significant to .75 with $p<.01$ for planning and a grade for Mechanics. The skills 'analysis' and 'planning' showed the largest number of significant values. Totalling all skills we found correlations with $p<.01$ for most physics grades and for PASS (.43) and SCORE (.46) for the university students, whereas the students of higher vocational education had correlation coefficients for PASS (.30) and SCORE (.47) with $p<0.05$.

Correlation coefficients for the skill measures are lower than those of leaving examination grades, and question is whether these measures offer any additional information relevant for predicting success in higher education. To investigate this we split the students according to grade into 'high', 'medium', and 'low' and repeated the analysis for each subgroup. The resulting groups are relatively small and results do not reach a confidence level higher than $p<.05$. At this level, there are indications that within these groups students with higher scores for problem solving skills have more success than those with low scores.

Discussion and Conclusions

This study has identified a number of problem solving skills, which play a central role in first year examinations of higher education, but do not get much attention in Dutch secondary education. As they form stumbling blocks for the students, they seem to get little attention in higher education as well. We consider the description of these skills as a description of tacit objectives of first year courses. The results also indicate that mastery of these problem solving skills at the beginning of higher education relates to success in the first examinations, even within groups of students with the same leaving examination grade.

We conclude that attention for a limited number of problem solving skills within the physics taught in the last forms of secondary education could help to streamline

the step from secondary to tertiary education in the scientific and technical sectors. The results offer a basis for further research on the possibility of teaching problem solving skills at this level. The description of specific skills and their qualities, and the specimen problems of this study will be used in further research on the teachability of a number of skills at secondary level. Future results, we hope, could offer guidelines for the development of didactic models, curricula, and study material for secondary schools.
For higher education we conclude that more attention for the 'tacit objectives' made explicit in this study could improve examination results in the first year. Some institutes for higher vocational education are already making efforts in this direction. The results of this study offers guidelines for this type of remedial courses. Universities have shown less explicit interest for problems of success in the first year. We hope that the results of this study could stimulate discussion and offer possibilities for further research within higher education, a necessary complement to studies at secondary level.

References

Alexander, P.A., & Judy, J.E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. Review of Educational Research, 58, 375-404.

Dyck, M. van; LICOR (1990). Studievaardigheden in het hoger onderwijs: Wat moet? Wat kan? [Study skills in higher education: What is necessary? What is possible?] Proceedings Landelijke dag studievaardigheden '90 (pp. 128-140). Technische Universiteit Eindhoven & Hogeshcool Eindhoven.

Jong, T. de, & Ferguson-Hessler, M.G.M. (1984). Strategiegebruik bij het oplossen van problemen in een semantisch rijk domein: Electriciteit en Magnetisme. [Using strategy in problem solving in a semantically rich domain: Electricity and Magnetism.] Tijdschrift voor Onderwijsresearch, 9, 3-16.

Jong, T. de, & Ferguson-Hessler, M.G.M. (1991). Knowledge of problem situations in physics: d comparison between good and poor novice problem solvers. Learning and Instruction, 1, 289-302.

Perkins, D.N. & Salomon, G. (1989). Are cognitive skills context-bound? Educational Researcher, 16-25.

Schoenfeld, A.H. (1985). Mathematical problem solving. Orlando: Academic Press, inc.

Taconis, T., Stevens, M. & Ferguson-Hessler, M.G.M. (1992). Probleemoplossingsvaardigheden in het natuurkundeonderwijs van de bovenbouw van het voortgezet onderwijs. [Problem solving skills in physics teaching in the final years of secondary education]. Eindverslag SVO project 0041. Vakgroep didactiek natuurkunde, TU Eindhoven.

247

ELERIAT: A FLEXIBLE EXPERIMENTAL LEARNING ENVIRONMENT FOR RESEARCH ON INDUCTION AND ANALOGICAL TRANSFER

Sigmar-Olaf Tergan, & Uwe Oestermeier, Department of Learning Research, German Institute for Distance Studies at the University of Tübingen, Tübingen, Germany

Abstract

According to theory and empirical research processes of inductive reasoning and analogical transfer are crucially dependend on critical aspects of the structure-mapping process underlying analogical transfer of knowledge from a source domain to a target domain (Gentner, 1983, 1989; Holland et al., 1986).

In this process, individual cognitive factors, for example, individual preconceptions, strategies in hypothesis testing, kind of learning mode (guided discovery learning, self regulated exploratory learning), surface vs. structure oriented processing seem to play an important role in impeding or facilitating successful inductive problem solving and analogical transfer (Gick & Holyoak, 1987; Vosniadou & Ortony, 1989).

Processes of inductive problem solving and analogical transfer are also dependend on some specific environmental and instructional variables. These variables are, for example, the similarity of successive problem situations in a problem solving environment with respect to surface as well as structural features, the kind of instructional means for facilitating access to potential source problems, for focussing attention to structural relations in the source domain, and for promoting schema induction (Gick & Holyoak, 1980, 1983, 1987; Holland et al., 1986; Gentner, 1989).

Modern cognitive research on inductive analogical reasoning so far has concentrated on processes relevant for solving a particular target problem under conditions of a well understood source problem (Gick & Holyoak, 1980, 1983; Holland et al., 1986). Relatively little is known about the potential of different instructional conditions for fostering induction of relevant structural relations inherent in the source problem (e.g. Gick & Holyoak, 1983; Vosniadou & Ortony, 1989). Even less is known about the possible potential of inductive analogical reasoning for fostering processes of knowledge restructuring under conditions of existing misconceptions about the structural relations inherent in the source problem.

The paper presented is an outline of the characteristic psychological, didactical and technological features of a computer-based flexible experimental learning environment in the domain of elementary mechanics ("superposition of motion"). The environment has been implemented on Macintosh using HyperCard and PASCAL. It has been developed to foster the design and performance of experiments in order to study assumptions concerning the effects of instructional variables on inductive analogical reasoning and knowledge restructuring.
At the present state of the development the system consists of 10 modules some of them being structured into submodules. The modules are (1) Content overview, (2) General introduction, (3) Outline of the problem situation, (4) Learning goals, (5) Analogical problem solving environment, (6) Playground for experiments, (7) Simulations, (8) Tutoring module, (9) Glossary, (10) Adjunct questions.
Modules 1 - 4 provide informations on the didactic structure of the system, the general problem situation in the domain "superposition of motion", and the central learning goal: the discovery of the structural relations inherent in the implemented subject-matter domain.

The modules 5 - 7 are the kernal modules of the experimental environment.
The "Analogical problem solving environment"- module provides a sequence of problem solving tasks. The sequence of the tasks can be fixed by the experimenter. The tasks are similar with respect to structural but different with respect to surface features. In a sub-module ("error diagnosis") a simple form of cognitive modelling is implemented. The module interpretes some issues of erroneous individual problem solutions and provides qualitative adaptive feedback in terms of indicating individual bugs and misconceptions possibly relevant for the incorrect problem solutions.
The module "Playground for experiments" is a virtual environment for exploratory learning. The learner when using the environment may plan and run her/his own experiments on "superposition of motion" and check the validity of the tested assumptions by inspecting the outcome of corresponding simulations.
The "Simulation"-module consists of three submodules, each of the modules providing dynamic visualization and highlighting different aspects of the structural relations inherent in the subject-matter domain.

Modules 8 - 10 provide additional informations for learning support.
The "Tutoring"-module provides different aspects of knowledge about the content domain. In the sub-module "physical principles" declarative knowledge about the domain-specific physical concepts is presented. The sub-module "Schema for problem solving" provides knowledge about the general procedure in solving problems on "superposition of motion". The "Glossary" contains informations about unfamiliar domain-specific terms.
The "Adjunct question"-module is made up of multiple-choice questions concerning problems on "superposition of motion". The questions serve two functions: They may be used for the assessment of domain-specific knowledge. They may also be used by the subjects for metacognitive control of individual progress in knowledge acquisition.
The sequence of the modules and sub-modules as well as informations on each Hypercard-card may be selected and fixed by the experimenter by means of a

"design-control card" according to her/his individual plan for experimental design, thus, generating a closed learning environment. The sequence may also be be made freely accessible to the learner by simple mouse-click of the experimenter, thus, creating an open learning environment to be used in an exploratory learning mode.

The presentation concentrates on the psychological and didactical rationale underlying the system. It outlines its modular HyperCard-based structure, its flexibility for designing different experimental environments on top of its modular structure, and its facilities for assessing qualitative and quantitative aspects of individual cognitive processes and performance data.

The results of an experiment on the effects of different sequencings of problem solving tasks and types of instructional support (e.g. simulation) on knowledge restructuring will be presented. The experiment had been performed on the basis of a preliminary version of the system and had contibuted considerably to the refinement of the system's rationale.

References

Gentner, D. (1983). Strcture-mapping: a theoretical framework for analogy. Cognitive Science, 7, 155-170.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.). Similarity and analogical reasoning (pp. 199-241). Cambridge, MA: Cambridge University Press.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. Cognitive Psychology, 12, 306-355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.

Gick, M. L., & Holyoak, K. J. (1987).The cognitive basis of knowledge transfer. In S. M. Cormier & J. D. Hagman (Eds.). Transfer of learning: contemporary research and applications. New York: Academic Press, Inc.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). Induction: Processes of inference, learning, and discovery. Cambridge, MA: MIT Press (Bradford Books).

Vosniadou, S., & Ortony, A. (1989). Similarity and analogical reasoning: a synthesis. In S. Vosniadou & A. Ortony (Eds.). Similarity and analogical reasoning (pp. 1-17). Cambridge, MA: Cambridge University Press.

BEST COPY AVAILABLE

SEMANTIC NETWORK THEORY APPLIED TO MATHEMATICS EDUCATION: A MICROANALYSIS

Gerhard Steiner, University of Basel, Basel, Switzerland

According to semantic memory theory knowledge of different kinds can be represented in semantic networks. The main feature of network-type represented knowledge is that conceptual nodes are interconnected by relations to each other representing - as a whole - the meaning of some superordinate concept.
The process of constructing semantic networks contains usually three different subprocesses:

- connecting conceptual elements by relations (in learning from text mainly by verbs,
- condensing large patterns of connected conceptual elements into higher order elements (usually nouns) in order to save processing capacity as well as storing space in working memory, and
- organizing the connected elements into a coherent pattern.

This process of constructing semantic networks takes place e.g. while observing the environment, but mainly in the process of listening to or reading some text information or in the process of integrating figural information from a text. Learning means

- either constructing new semantic networks or
- restructuring (or adapting) already existing ones either by elaborating or by reducing them according to the purpose of the learning activity.

Retrieving information from memory is achieved

- by accessing conceptual nodes and following the connecting relations (Collins & Quillian, 1969),
- by unfolding formerly condensed conceptual higher order nodes or
- by reconstructing parts of the networks starting from known elements.

Besides the large number of semantic networks that represent the individual knowledge of a person (semantic and episodic, (Tulving, 1972)), special kinds of networks exist, among others numeric networks or micro-networks that are constructed early in the scholastic career of a child while learning arithmetics (Steiner, 1983, 1988). Contrary to common semantic networks the conceptual elements of a numeric micro-network are numbers, and the relations between them are mathematical operations, thus, the number of relations being restricted to just four at the beginning, rising up to a dozen or two when all mathematical operations are grown in the child's (or the adult's) mind several years later. Some elementary examples will show what numeric micro-networks look like in their

earliest form. There are also algebraic-mathematical micro-networks (AMMNs), which are the main topic in what follows. When talking about algebraic-mathematical schemata I always mean "active parts of networks", following Norman and Rumelhart's interpretation (Rumelhart & Norman, 1973, 1976; Rumelhart, 1978).

The educational starting point for my dealing with mathematical learning processes here is that startling fact that for many Senior High-School students mathematics learning or mathematical thinking is - at least within several topics such as elementary algebra, transforming terms, dealing with fractions or factorizing - nothing else than manipulating mathematical symbols, often without the faintest understanding of what they are doing, instead of an appropriate reasoning within well organized AMMNs. By our "progressive transformation" approach we aim at inducing a new culture of mathematical reasoning in our students.

In a micro-analysis I shall go into the details of the mathematical reasoning and learning processes in solving "worked examples" as the Australian group of Sweller and Cooper (1985) sees it as well as into the details of Zhu and Simon's (1987) constructions of production systems for factorizing trinomials in algebra learning.

The Sweller and Cooper (1985) approach stresses the fact that students should not waste their cognitive capacities for the search of possible problem solving procedures but, instead, focus on the schemata that are adequate and necessary to solve a problem. This focus is made possible by means of going through "worked examples", i.e. problems that are already solved and written down and can be analyzed step by step by every student at his/her own pace. Attention is directed towards the single states of the problem solution and the moves that connect these states. From this procedure students learn to derive the governing rules for solving the particular problem. The whole procedure does not ensure that constructed schemata or structures can be transferred into new problems - at least novices do have troubles in doing so.

Similarly, Zhu and Simon (1987) use "worked examples", at least in one of their experimental groups. Their students solved factorizing tasks without support by a teacher by learning "production systems" (we would say "schemata") to find for a trinomial like e.g. x^2+7x+6 a pair of factors leading to "6" in the constant term and adding up at the same time to "7" in the linear term. Central to our empirical work is the goal to stimulate in the students' mathematical learning processes transformational procedures in tackling algebra problems (such as rewriting fractions or factorizing or working with functions and their figural representations). We use a progressive transformation approach starting from already constructed AMMNs. A progressive transformation is a very close transformation of an AMMN that avoids far structural transfer and - by doing so - warrants an "optimal fit" as far as task difficulty is concerned which in turn provides the student with a sense of self- efficacy in that the current step in network construction is large enough to be a challenge for the student but short enough to prevent him/her of loosing sight of the inner structure of the problem.

Here is an example: The students have calculated $(x+2)(x+3)$ and obtained the result of x^2+5x+6. The "semantic space" of the equation $(x+2)(x+3)= x^2+5x+6$ is constituted

- by the adequate multiplications and the addition that transform the left hand side of the equation to what is stated on the right hand side; thus, its meaning lies in the activities of executing the mathematical operations. Neither Sweller and Cooper (1985) nor Zhu and Simon (1987) go into the meaning problem, i.e. into the network construction problem.
- Moreover, the meaning of the equation can also be represented in a drawn pictorial configuration (in our example by means of a rectangle), and
- it could be represented in the form of a function's graph if transformed into $y=x^2+5x+6$, all these representations being part of the "semantic space" or the AMMN, respectively, of that equation which is the base for the first progressive transformation (our instructional approach) that is expressed by the following question: What will happen to the right hand side of the equation if we transform the left hand side into $(x+1)(x+3)$ [instead of $(x+2)$]? The student must not just multiplicate the two binomials (in the brackets); instead, he/she is supposed to anticipate in face of the already known result of the first equation what effect(s) the transformation will produce. Post hoc only, i.e. when the anticipation (a kind of a hypothesis) is liably stated by the student he/she may calculate the multiplication to check the correctness of the anticipation. Then the progressive transformation goes on: The actual state of the equation is, again, the starting point for the next progressive step: Transform $(x+1)(x+3)$ into $(x+1)(x+6)$! The student is led from one transformation - including a compulsory anticipation and its control - to the next one, some of them being rather smooth and gradual (e.g. different integers), some rather essential (a minus sign instead of a plus sign) and more difficult to cope with.

While Sweller and Cooper (1985) aim at a construction of single schemata, and Zhu and Simon (1987) lead students to detect the production systems that are relevant in factorizing tasks, we focus on the construction of more encompassing as well as coherent AMMNs. By doing so, we concentrate not just on schemata construction but on the functions of schemata, on their usefulness and transferability; there is not only a shift of attention from solution search techniques to the relations between parts of the problem states (what Sweller and Cooper successfully do) but a shift towards the generation of such relations which constitute AMMNs.

It seems to me that Sweller and Cooper (1985) as well as Zhu and Simon (1987) strive after short-term goals: Students should learn the rules for either algebraic term transformations or factorizations. "Worked examples" are presented, schemata or production systems should be acquired and - in the next step - applied to similar tasks. Our goal is quite different: We aim at long-term changes: the acquisition of the knowledge about how to construct AMMNs by progressive transformations of available AMMNs, to provide the student with the persuasion that the operations or the problem solving steps are feasible and to have him/her experience at large his/her self-efficacy in coping with algebraic tasks.

Results from a pilot study are presented in which the progressive transformation approach has been tested against a conservative-traditional mathematics teaching and learning approach. Different measures have been used: motivational ones, assessments of task difficulty, of the security about the correctness of the particular test results, thinking aloud protocols of problem solving processes, and different error analyses. As far as systematic errors are concerned the results are promising

in favour of the progressive transformation approach. Theoretical as well as methodological problems in regard to AMMNs, their representation, assessment, and use will be discussed.

Keywords
Semantic network theory, construction of algebraic-mathematical micro-networks (AMMNs), progressive transformation, self-efficacy (motivation), "new culture of mathematical reasoning", thinking aloud protocols, error analysis. \

References

Collins, A.M., & Quillian, M.R. (1969). Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 8, 240-247.

Rumelhart, D.E. (1978). Schemata: The building blocks of cognition:. In R. Spiro, B. Bruce & W. Brewer (Eds.), Theoretical issues in reading comprehension. Hillsdale, NJ: Erlbaum.

Rumelhart, D.E., & Norman, D.A. (1973). Active semantic networks as a model of human memory. Proceedings of the Third International Joint Conference on Artificial Intelligence. Stanford, California.

Rumelhart, D.E., & Norman, D.A. (1976). Accretion, tuning, and restructuring: three modes of learning. San Diego: University of California, Technical Report no. 63. Steiner, G. (1983). Number learning and constructing coherent networks by using Piaget-derived operative principles. In M. Zweng, T. Green, J. Kilpatrick, H. Pollak & M. Sydam (Eds.), ICME Proceedings of the Fourth International Congress on Mathematical Education. Boston: Birkhauser.

Steiner, G. (1988). Lernen. Zwanzig Szenarien aus dem Alltag. Bern: Huber.

Sweller, J., & Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. Cognition & Instruction, 2, 58-89.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), Organization of Memory. New York: Academic Press.

Zhu, X., & Simon, H.A. (1987). Learning mathematics from examples and by doing. Cognition & Instruction, 4, 137-166.

THE USE OF ROLE PLAYING IN INTERPERSONAL SKILL TRAINING

Geralien A. Holsbrink-Engels, Department of Education, University of Twente, Enschede, The Netherlands

Introduction

In many outlines of interpersonal skill training programs, role playing is used as an instructional method. From the field of interpersonal skill training is not much known of the use of role playing. Many authors (Shaftel & Shaftel, 1967, 1982; Joyce & Weil, 1980, Tistaert, 1987; Nijkerk, 1988) only give (global) description of this instructional method in an optimistic and positive way without empirically confirmation. Further, the description of role playing concerns in most cases regular education. Not much is known over the use of role playing in corporate training. The main purpose of this study is to draw together information on the use of role playing in corporate training, in particular interpersonal skill training.

Method

Subjects

The population of the study would be corporate trainers.Variables controlled by exclusion are (1) the trainer must have experience with role playing in interpersonal skill training, and (2) the trainer must have held position for a minimum of one year.

Corporate trainers (N=200) are randomly selected from the registers of the Dutch Association of Corporate Trainers (Nederlandse Vereniging van Opleidingsfunctionarissen) and from the registers of the VETRON (Vereniging van Trainings- en Opleidingsinstituten in Nederland).

Materials

A written questionnaire consisting of about 60 items will be used to measure facts and attitudes toward the use of role playing in corporate training. The questions and statements refer to (1) the use of role playing and (2) disadvantages of and problems with role playing. Based on literature review, the use of role playing will be specified into the four training stages of role playing: (a) the introduction (objectives, method and procedure, context, props or additional material & warming up), (b) instruction (choice and instruction of players, instruction of observers), (c) run session (start the actual role-play, interruption, provide information, control time and stop) and (d) debrief (clear up role play, draw conclusion and draw general conclusions). A specification of disadvantages and problems with role playing will be enumerated by Van Ments (1989) and Georges

(1990). Two different structured response mode will be used: Checklist response and a scaled response (Tuckman, 1988).

Procedure

A review of literature in the area of role playing is undertaken to identify the variables to be measured. These variables serve as starting point for the construction of the questionnaire. In a pilot study, the questionnaire will be tested for formulation and styling of the questions and statements. Revision will take place on the results of the pilot study.

The initial mailing of questionnaire to a sample of respondents includes a cover letter, the questionnaire itself and a stamped, return-addressed envelope. After a period of about two weeks to a month has elapsed, a second mailing will start to correspond with the nonrespondents. The second mailing will include another letter soliciting cooperation, another questionnaire and another stamped return-addressed envelope in case the respondent can not find the original ones. If the second mailing fails, additional mailings and telephone calls will be employed to elevate the return. This study will be striving for a return of more than 75%.

Results, Conclusion & Discussion

At the beginning of 1992, the data will be collected. The checklist responses and rating scales will be precoded. Each response can be directly converted into a score. The scaled responses can be considered to be interval data. The data lend themselves to frequency and chi-square analysis. A reliability check for internal consistency (i.e. the Spearman-Brown Prophesy Formula, Nunnally, 1978) will be made.

Reference

Georges, J.C. (1989). The Hard Reality of Soft-skills Training. Personnel Journal, 4, 41-45.

Joyce, B.R., & Weil, M. (1980). Models of teaching. Englewood Cliffs, W.J.: Prentice Hall.

Nijkerk, K.J. (1988). Het rollenspel. In: Kessels, J.W.M. & Smit, C.A. Handboek Opleiders in Organisaties [Handbook for trainers in organizations]. Deventer: Kluwer Bedrijfswetenschappen.

Nunnally, J.C.(1978). Psychometric Theory. New York: McGraw-Hill.

Shaftel, F., & Shaftel, G. (1982). Role-playing in The Curriculum. Englewood Cliffs W.J.: Prentice-Hall.

Tistaert, G. (1987). Rollenspel. In het losbladig Onderwijskundig lexicon [Educational lexicon]. Alphen aan de Rijn: Samson.

Tuckman, B.W. (1988). Conducting Educational Research. New York: Harcourt Brace Jovanovich.

Van Ments, M. (1989). The effective use of role play: a handbook for teachers and trainers. London: Kogan Page LTD.

EFFECTS OF TRAINING IN ANAMNESTIC INTERVIEWING SKILLS

S.M. Bögels, & T.G.C. van Mourik, Department of Educational Development and Research, University of Limburg, Maastricht, The Netherlands

Introduction

The anamnestic interview is the basic clinical medium in mental health care. Interviewing and problem solving are important professional skills for mental health care workers like doctors and psychologists. Therefore, training in these skills are part of most health care education programs. Although much energy is spent in developing training programs, controlled studies on the effects of these training programs are lacking.

Studies on the effectiveness of anamnestic skill training programs show many methodologic weaknesses. Most studies use "semi-controlled" designs, like cross-sectional designs (Kraan et al., 1990) or pre-posttest designs without a control group (Gask et al., 1987). Moreover, evaluation studies do not account for inter-case reliability of the skills (i.e. consistency of performance across cases). The research on medical (interviewing) skills has consistently indicated that examinee performance on one case is a poor predictor of performance on other cases (Van der Vleuten & Swanson, 1990). Research on psychiatric interviewing skills has revealed that a minimum of four interviews is necessary to obtain reproducible scores of psychiatric interviewing ability (Bögels, Schmidt & Van der Vleuten, 1990). Effect studies using only one interview to measure progress in interviewing skills, may be of limited importance. Effect studies mostly use direct observation of behavior as a dependant variable. Clearly, observational measures are the most important effect measures in evaluation studies. Yet, using *only* observational measures to determine training effects can lead to mechanical behavior of trainees (Stone, 1982). Evaluation research needs to focus on various training effects: behavior change, cognitive change, and experiences of simulated patients.

The purpose of the present study was to evaluate, in a controlled way, effects on students of a short training in anamnestic interviewing skills for mental health problems. We intended to answer the following questions:

- Do anamnestic interviewing skills improve as a result of training?
- Do cognitive skills (recall of relevant patient information and problem solving) improve as a result of training?
- Does patient satisfaction improve as a result of training?

Method

Subjects

Sixteen (of 60) second-year students in the Mental Health Sciences of the University of Limburg voluntarily participated in the study.

Cases

Twelve different case histories were developed, derived from real patients. The cases concerned patients with affective and psychotic disorders. The cases were enacted by twelve experienced simulated patients, partly psychiatric nurses.

Design

A pre-pre-post, or time-series design, was used to measure progress in anamnestic interviewing skills. Three weeks before the training students were tested for the first time (pretest 1). Prior to the training this test was repeated (pretest 2). After the three-week period in which students were trained, they were tested again (posttest).

Training

The training 'Anamnesis' consists of six group meetings of three hours, and homework assignments. The objective of the training is to learn anamnestic interviewing skills and to make correct diagnosis and treatment planning, based on data gathered during the interview. By means of video demonstrations, roleplay, and feedback, students learn psychiatric interviewing techniques.

Assessment

- Anamnestic Interviewing Scoring list for Psychiatric problems (AISP).
  The AISP (Bögels, Schmidt & Van der Vleuten, 1991) consists of two scales called "Process Skills" and "Content Skills". Process scores reflect ratings on ten aspects of interviewing technique. Content scores reflect rating on thirteen aspects of information gathering technique.
- Patient Satisfaction with Communication Checklist (PSCC).
  The PSCC (Kraan & Crijnen, 1987) consists of eighteen items reflecting empathic understanding, and clearity of communication of the interviewer. On each item, patients indicate how strong they agree or disagree.
- Recall.
  Audiotaped recall sessions were organized to obtain recalled patient information from the subjects. Scoring lists were developed to measure the amount of relevant patient information recalled. The lists were based on the twelve case histories roleplayed by the simulated patients. Each list contained 30-40 items of relevant patient information. Five categories were distinguished: diagnosis, differential diagnosis, aetiologic factors, predisposing factors, question for help & previous therapy.
- Problem solving.
  Problem solving quality was measured by means of a semi-structured questionnaire, validated in previous research on anamnestic interviewing skills (Kraan & Crijnen, 1987). The interviewer responded to six open-ended questions about diagnosis, explanatory hypotheses and hypotheses about further management.

Procedure

During each test, subjects were instructed to interview four simulated patients, each for a period of 30 minutes. The interviews were videotaped. Immediately after each interview students were asked to report all patient information. Between the interview and the recall procedure subjects had to do a half-minute counting task, in order to control for recency effects (Peterson & Peterson, 1959). Finally, they completed the problem solving questionnaire. The simulated patients simultaneously completed the PSCC.

Data analysis

Four graduate students Mental Health Sciences were asked to rate the skills of the interviewers. All raters were at the end of their studies and had completed their six-month term of probation in a mental health institute. Raters were extensively trained in using the AISP. Because prior research (Bögels, Schmidt & Van der Vleuten, 1990) showed that one rater provided sufficient reliable ratings, the videotaped interviews were scored by only one rater. Raters were nested within patients; rater 1 scored the interviews of all subjects with patient A, rater 2 scored the interviews of all subjects with patient B, and so on. All raters judged the interviews with one patient on pretest 1, pretest 2 and posttest. Raters were blind to whether the interviews were pre or post training. Audiotaped recall of relevant patient information was rated by a graduate student Mental Health Sciences. In order to judge the answers of the students on the problem solving questionnaire, a scoring key was developed based on the answers of five experts. Another graduate student Mental Health Scientist rated the completed problem solving questionnaires of the students, based on the scoring key.

Statistical analysis

Inter-rater reliability was estimated by means of correlational analysis. Generalizability analysis was used to estimate overall reliability of the test (Cronbach et al., 1972). Multivariate analysis of variance was used to measure effects of training.

Results

Inter-rater reliability of Process and Content Skills, based on the ratings of all four raters of ten subjects, was sufficient. The pearson correlation coefficient for Process Skills was .84, for Content Skills .87. Preliminary analyses of the effects of training show no differences between pretest 1 and pretest 2, and significant differences between pretest 2 and posttest for Content Skills and recall of relevant patient information. Final results will be presented at the conference.

References

Bögels, S.M., Schmidt, H.G., & Van der Vleuten, C.P.M. (1990). Ontwikkeling van een vaardigheidstest voor het anamnestisch interview in de geestelijke gezondheidszorg. In:J. Hoogstraten & W.J. van der Linden (red.) Methodologie. Amsterdam: Stichting Centrum voor Onderwijsonderzoek.

Cronbach, L.J., Goldine, C.G., Nanda, H., & Rajaratnam, N. (1972). The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles. New York: Wiley & Sons.

Gask, L., McGrath, D., Golberg, D., & Millar, T. (1987). Improving psychiatric skills of established general practioners. Medical Education, 21, 362-368.

Kraan, H.F., & Crijnen, A.A.M. (1987). The Maastricht History-taking and Advice Checklist: Studies of instrumental utility. Proefschrift, Amsterdam: Velden van den Hazelaar.

Kraan, H.F., Crijnen, A.A.M., De Vries, M.W., Zuidweg, J., Imbos, T., & Van der Vleuten, C.P.M. To what extent are medical interviewing skills teachable? Medical Teacher, 3/4, 315-328.

Peterson, L.R., & Peterson, M.J. (1959). Short term retention of individual verbal items. Journal of Experimental Psychology, 3, 193-198.

Stone, G.L. (1982). Evaluating the effectiveness of skills training programs. In: E.K. Marshall,

P.D. Kurtz & associates (red.) Interpersonal helping skills. San Francisco: Jossey Bass Publishers.

Vleuten, C.P.M. van der, & Swanson D.B. (1990). Assessment of clinical skills with standardized patients: state of the art. Teaching and Learning in Medicine, 2, 58-76.

260

LEARNING STYLES OF LEARNERS AT PRIMARY SCHOOLS: THEORETICAL AND EMPIRICAL PROJECT

Hana Krykorková, Czechoslovakia

Introduction

The goal of this paper is to inform:

- about the global project goal
- about the empirical goal and some investigation results
- about the development and improving goals

Main present theoretical question is considered as the mapping relationship between two principal components of learning style- predominantly innate component and situationally conditioned component.

The part of this project is to assess a population of czechoslovac learners of primary school as well; one foreign questionnaire of learning styles will be standardized including factor analysis and czechoslovac norms/questionnaire of learning styles from Entwistle is determined and will be modified for learners of primary school-with proper permission of the author/.

Measurement of relation will be provided between the selected dimensions of cognitive style in primary learners/reflexivity-impulsivity, dependance on the field, memory dimensions of cognitive style/ on one hand, and between the dimensions of their learning styles and scholar achievements, on the other, repectively. Also an operational structure of learning tasks for their resolution by learners is to be taken into account.

First measurement of relations was provided between the reflexion-impulsivity dimension of cognitive style/ MFFT/ and between scholar achievement. On the whole we may say that the research data have drawn attention to the significance of reflexiveness and impulsiveness in children, to the weighty role of this trait from the aspect of effectiveness of the teaching activity and its relation to school achievements.

On the basis obtained data will be elaborated advices and recommendations for schools and teachers how to implicate the data on learning-teaching styles.

KNOWLEDGE INTEGRATION BY MEANS OF OPTIMIZATION- AND TROUBLESHOOTING-TASK WITH COMPUTER SIMULATION

Frans Pijnenborgh, Faculty of Philosophy and Social Sciences, Educational Research Group, Eindhoven University of Technology, Eindhoven, The Netherlands

Introduction to the problem

The main objective of the project is to encourage students to integrate several subdomains within the domain of chemical process technology. For this we use, as a basis, a computer simulation program, ChemCAD, on chemical process technology. Learning with computer simulations is considered important in education because it may engage learners in exploratory learning. As a consequence it enables students to construct actively a knowledge base. To ensure that in this construction an integration of subdomains will take place it is necessary to provide additional tasks to the simulation.
In this study the influence of two kind of tasks, *optimization* and *trouble shooting*, on knowledge integration and exploratory learning was investigated. In particularly the focus was on the learning strategies used in the different tasks and differences between experts and novices. Based on the results obtained we will state some conclusions on the problems that arise in exploratory learning and we will make suggestions on the kind of support and stimulation is needed.

Method

Domain

The domain involved was chemical process technology. The subjects worked with ChemCAD: a professional interactive computer simulation program for engineers in the chemical industry. Among other things ChemCAD offers the possibility of simulating (complex) chemical plants. Two such plants were simulated. Subjects had to optimize one plant and trouble shoot the other.

Subjects

The experiment was conducted with 19 pairs of third and fourth-year students of chemistry and three experts of the department Chemical Process Technology at the Eindhoven University of Technology. The experiment was conducted in the normal practical work and pairs of subjects were divided almost evenly over four sessions. The experts participated in the experiment at an individual basis.

262

Conditions

Two different tasks, optimization and trouble shooting, were offered to the subjects. Ten out of nineteen pairs performed the trouble shooting task on the first day followed by the optimization task on the second day and nine out of nineteen pairs followed the opposite order. In the optimization task subjects had to optimize an already working (simulated) plant and could, if necessary, add or delete equipments to the plant. In the trouble shooting task subjects were asked to solve the problems in a plant, which didn't work any more after a maintenance service. In the trouble shooting task it was not allowed to add or delete equipments.
Our prediction was that experts would perform better then students in the optimization task because of the greater domain knowledge of the experts. (With a good performance we mean that little trials are needed to solve the task and that most of the variable changes belonging to those trials are based on hypothesis.) On the other hand we did not expect a better performance of the experts in the trouble shooting task because their practical experience and expectations are of less use in a artificial (simulation) environment then in a normal situation. In a simulated environment it is possible to build in some faults which are not common in normal situations. Also, it was expected that the experts would have a lower 'fault-score' than the students on the pre- post-test of both tasks, the optimization in particularly. We also assumed that there was a learning effect for the second day independently of the order of presenting the tasks.

Procedure

The subjects worked with the simulation for three hours on both days. Before they started to work with the simulation the had to fill in a pre-test, in which we asked the subjects to assess the equipments present according to the level of attention they thought they would give these equipments after reading the task. The subjects could choose from three categories, 'large', 'normal' and 'little' attention. After they had finished the task they were asked to fill in a post-test at which they were asked to give the actual amount of attention which they thought each of the equipments deserved.
For observing their learning process we asked the subjects to fill in a dedicated check-list after each simulation run. This list consisted of questions directed at the previous run, the current run and the run to come. Some examples are 'where did you make a variable/parameter change?', 'what parameter/variable did you change?', was the change based on an expectation or just a feeling? and 'if you had an expectation did it come true?

Results

Data are still being analyzed but first preliminary results indicate that the experts perform better on the optimization task than students do. In fact, experts use significant less trials on the optimization task then students do. Also most of the variable changes belonging to those trials are based on hypothesis. As a result of the lower number of trials experts take longer time between two trials. In the trouble shooting task on the other hand there is no difference between the experts and the students, they both need the same amount of trials and the time between two trials is more or less the same. Experts and students use different strategies on the optimization task, experiments performed by expert are more hypothesis driven

whereas those of the students are more based on selective trial & error. For the trouble shooting task there is some kind of resemblance between both subject groups, they both make use of selective trial & error. At the pre and post test we can see a decrease in 'fault scores' on the second day. This is especially true for the newly introduced equipment. For each of the tasks the total 'fault scores' on the first day were compared with the total 'fault scores' on the second day. For the optimization task we did find a tendency of lesser 'fault scores' on the second day and for the troubleshooting task we did find a slight tendency in favour for the first day. A closer look at the individual equipments shows that the equipments contributing to the solution have a significant or almost significant lower 'fault score' on the second day.
In regard with the order of presentation of the different tasks no effect was found.

Conclusions and discussion
At the moment we are not ready to state firm conclusions because the data are not fully analyzed. There are however strong indications that there are differences between experts and students in the optimization task but not on the trouble shooting task. One of the explanations may be that experts can not use their greater knowledge on tasks where they can not make use of their practical experience and expectations. On the other hand we have some evidence that offering tasks in an exploratory environment can help enlarge domain knowledge and can accomplish knowledge integration. This can be illustrated by the shift from a selective trial & error strategy used by students to a more hypothesis driven strategy. Also we see a little decrease in the number of trials if we compare one type of task on the first day and second day.
Although we did not find an overall effect in regard with the order of presentation there still might be a local effect of order, in this case a negative effect. We think that this local effect can be seen in the presentation order optimization followed by troubleshooting.
There is strong indication that in the troubleshooting task on the second day the first couple of changes are based on successful changes in the optimization task on the first day. This might explain the slight tendency in favour of the first day for the troubleshooting task. However these data need further analysis.

THE EFFECT OF GRADE AND LOGICAL REASONING ABILITY ON CONCEPTIONS RELATED TO SELECTED SCIENCE CONCEPTS OF FIFTH, EIGHTH AND ELEVENTH GRADE STUDENTS

Aysenur Yontar, Turkey

In the interactions with school children who are learning science, it is natural to observe some individual differences in their ability to understand science concepts, conduct investigations and solve some specific problems. While some students are capable, others can demonstrate inappropriate reasoning patterns. Many students find science subjects difficult to learn and they can not develop sound understanding about related concepts.

Majority of science concepts studied at junior and senior high school levels are abstract. All Turkish children begin their study of science during fourth grade of primary school. Even in that level some abstract concepts take place in the science courses. Some students find those concepts difficult to learn. Logical reasoning abilities of students can explain some of those difficulties in learning. Such kind of concepts require more than performing logical operations on concrete reality. They require hypothetical deductive reasoning about unobservable entities. That kind of reasoning pattern is the characteristic of formal but not concrete operational reasoning. Related research studies indicate that 40%-75% of secodary school students have fail to reach the level of formal thought (Lawson and Renner,1975). If concrete thinkers can not develop sound understanding of abstract concepts because of their logical reasoning abilities then a large portion of secondary school students will find most of the science concepts difficult to learn.

Also variations in prior knowledge about the science concepts may cause different conceptualizations. Research studies indicate that, students have ideas about most of the science concepts before having formal science classes. So while learning science concepts, logical reasoning abilities and prior experiences about the concepts may shape the conceptions of students.

The present sdudy attempts to provide some information about conceptualizations of students in differrent grade levels and also it aims at putting forward the relationship of conceptualizations of students with their logical reasoning abilities.So, the purpose of the study is to investigate the effects of logical reasoning ability and grade level on conceptions of students, related to selected science concepts of fifth, eighth and eleventh grade students. Two science concepts are selected in order to conduct investigation. These are *particular*

nature of matter and *density*, which are first introduced students in the fourth and fifth grade science classes.

Sample of the study was choosen from nine public schools in Ankara. All of the schools were similar in their SES as well as their level of performance in terms of the results of Second Stage of the University Enterance Examination.

Due to the limitations of the study it was impossible to conduct a longitudinal study in order to understand the changes in conceptualizations of students by time exposure to schooling. So the researcher used stratified random selection while choosing the subjects from each grade level and 85 fifth, 73 eighth and 71 eleventh graders - N=229 - served as the sample of the study.

The data were collected through three diferent instruments; First instrument was the Turkish version of *Logical Reasoning Test*, originally prepared by G. Burney in 1974 (Cited in Sund, 1976) and subsequently translated into Turkish by H. Garibagaoglu (1985). This instrument was used in order to measure the logical reasoning abilities of subjects as stated by Piaget.

Because, one of the aim of the study is to investigate the structure of pupils conceptualizations, researcher felt that a paper pencil test was not sensitive enough for this purpose. So the researcher chose to use structured interview technique in order to obtain data about conceptions of subjects. *Tasks Related to Particular Nature of Matter* and *Tasks Related to Density* were the other two instruments developed by the researcher for measuring conceptualizations of the subjects related to selected science concepts. Both of the instruments were conducted, revised and finalized by using two pilot samples which constituted totally 93 pupils from different grade levels. *Tasks Related to Particular Nature of Matter* was an interview instrument which was used to measure the conceptions of subjects related to the concept of particular nature of matter and totally six tasks were included. *Tasks Related to Density*, was the other interview instrument, which included four tasks and used in order to measure the conceptualizations of subjects about the concept of density. All of the instruments were individually administered. Interview sessions were audiotaped and transcribed by the researcher.

The data were analysed by using *multi-dimensional contingency tables*. The study was concerned with three classificatory variables - grade, logical reasoning ability and conceptions -. A fitting model called *log-linear model* was utilized in order to analyse the data.

The results of the study are as follows; The effect of interaction between grade and logical reasoning ability on conceptions related to the selected science concepts as measured by *Tasks Related to Particular Nature of Matter* and *Tasks Related to Density* were statistically significant (p=.001).

The main effect of grade on conceptions related to selected science concepts are found to be statistically non significant at .05 significance level, while the main effect of logical reasoning ability on conceptions related to selected science concepts was statistically significant (p= .05).

The most important finding of the study seems to indicate that the effect of logical reasoning ability on conceptions of students is statistically significant. The results of the study did not indicate the importance of grade levels -age- in improving conceptions of students. Also analysis of the data indicates the alternative

conceptions of students about selected science concepts and show some tendencies among these conceptions.

References

Garibagaoglu, H. (1985). The relationship between cognitive development level and sex, age, science achievement and mathematics achievement.. Unpublished master's thesis, Middle East Technical University, Ankara.

Lawson, A., & Renner, J. (1975). Relationships of science subject matter and developmental levels of learners. Journal of Research in Science Teaching, 12(4), 347-358.

Sund, R.B. (1976). Piaget for educators. Colombus, Ohio: Charles E. Merrill Publishing Company.

THE EFFECTS OF A CONSTRUCTIVIST TEACHING STRATEGY ON USING LAY-IDEAS ON IONIZING RADIATION AND THE ABILITY TO MAKE RISK ASSESSMENTS

H.M.C. Eijkelhof, A.E. van der Valk, & P.L. Lijnse, Centre for Science and Mathematics Education, University of Utrecht, Utrecht, The Netherlands

Earlier studies have shown that a number of lay-ideas on radioactivity and ionizing radiation are very common among senior high school students. These lay-ideas have also been found in many newspaper reports in different countries and are recognized as important by radiation experts working in a variety of professions. Such ideas have appeared to be quite resistant to current ways of teaching the topic in physics classes. This does not only apply to traditional (fairly academic) teaching strategies but also to teaching this topic in the context of risk assessment, such as used in the Dutch PLON-unit Ionizing Radiation. It has been shown that some of these ideas make it difficult to judge the safety of situations in which ionizing radiation is used. Judgments of this kind are required according to the new Dutch physics examination programmes for senior high school (HAVO and VWO).

In this paper a study will be reported in which the effects of using a constructivist strategy of teaching the topic has been evaluated. Main characteristics of the strategy are the incorporation of activities in which lay-ideas are activated and discussed, and the attention given to the meaning of concepts and to their mutual relations.

In the study four hypotheses have been tested. They deal with the following issues:

- the productive role of lay-ideas in concept development;
- the mastering of concepts which are essential for risk assessment;
- the comparison of mastering these concepts between the experimental group and a control group;
- the comparison of using these concepts in making risk assessments between the experimental group and a control group.

Seven classes have been taught in the new way (i.e. the experimental group). Pre- and post-questionnaires have been completed by the students in which the use of lay-ideas, the mastering of concepts and the ability to make risk assessments have been evaluated. In addition to this, class discussions have been recorded and teachers have been interviewed. Both questionnaires have also been completed by students of a control group of seven classes in which the topic ionizing radiation

was taught in a traditional way. Finally a retention questionnaire has been completed by the students of the experimental group after about six months.
The results available so far show that the students of the experimental and the control groups do not differ in the use of lay-ideas before the topic is taught. However, after education on the topic, students of the experimental group use far less lay-ideas, have reached a much higher level of mastering the concepts and are much more able to make risk assessments for a variety of situations in which ionizing radiation is applied. The results of the retention questionnaire will be available early in 1992.
Analysis of the recorded lessons show that the constructivist activities generally lead to productive discussions, although some teachers have difficulties with applying the strategy in the way it was intended.
In the discussion part of the paper these results will be evaluated. Particularly attention will be given to the value of the various components of the strategy for teaching ionizing radiation and other topics in science education, and to problems of implementing this strategy.

THE CONTRIBUTION OF TREATMENT AND LOGICAL THINKING ABILITY ON STUDENTS' LEARNING OUTCOMES AT HIGH SCHOOL LEVEL

Petek Askar, & Omer Geban, Faculty of Education, Middle East Technical University, Ankara, Turkey

Introduction

One of the important goals of science education has always been to develop cognitive learning outcomes. These learning outcomes are intellectual abilities and recognition of knowledge in science. During the recent years research in science education has stressed the importance of cognitive factors that influence science achievement. Logical reasoning is one type of cognitive ability.

Several studies reported that there has been a significant positive relationship between cognitive structure and science achievement at high school level (Bender and Milakofsky, 1982; Herron,1975).The students in high school science courses are at a variety of developmental stages. Some of them have been at the concrete operational level, some formal and others in a transition stage. Some studies showed that there have been positive relationship between students' formal reasoning ability and science achievement (Chandran, Treagust and Tobin,1987).

Another important type of cognitive ability is science process skills. Process skill learning has become an important part of science curricula.

Today current curriculum organizations emphasize on problem solving strategy. If students are exposed to problem solving experiences then certain changes can occur in their cognitive structures. Some studies indicated that the problem solving approach has improved science achievement (e.g.Mulpo and Fowler, 1987; Chiappetta and Russell,1983). Also, in the study of Tobin and Capie (1982), this approach has produced greater intellectual abilities than the traditional approach. In summary, some factors in learning conditions such as logical thinking ability and treatment are important for influencing achievement in science and intellectual development.

The purpose of this study was to analyze the relationship among the independent variables , the treatment, logical thinking ability and the interaction between them and the dependent variables of chemistry achievement and science process skills.

Method

Subjects

The subjects of this study included 140 ninth grade students in a high school in Turkey. Four classes taught by two chemistry teachers were randomly selected. Each teacher had two classes. Each of two teaching approaches (problem solving vs traditional) was randomly assigned to one class of each teacher. Each class consisted of 35 students. The data was taken from 70 students in problem solving treatment, and 70 in the traditional treatment.

Instruments

Three instruments were used in this study. These were Chemistry Achievement Test, Logical Thinking Ability Test and Science Process Skill Test.

Chemistry Achievement Test was developed by Geban and Ozkan. The topics in this test were the mole concept, chemical reactions, gases, and solutions. The test was investigated by a group of experts in chemistry and science education and the course teachers. It includes 48 items. The alpha reliability coefficient was found as 0.88.

Logical Thinking Ability Test of Tobin and Capie (1981) was translated and adapted by the researchers into Turkish. It is a 10-item test and measures the abilities that are identifying and controlling variables, proportional, correlational, probabilistic, and combinatorial reasoning. The reliability coefficient was found as 0.77 for this study.

Science Process Skill Test of Okey, Wise, and Burns (1982) was translated and adapted into Turkish by the researchers. It includes 36 items and measures the abilities that are identifying variables, identifying and stating hypotheses operationally, defining, designing investigations and graphing and interpreting data. The reliability coefficient was found as 0.81 for this study.

Procedure

This study was conducted over approximately 9 weeks. Each instructional approach included the classroom instruction and the laboratory studies. The laboratory activities were taught by Geban. The classroom instruction was given by the classroom teachers. Lecture and discussion methods were used in the classrooms.

The traditional and problem solving group participates in six laboratory activities related to mole concept chemical reactions, gases, and solutions.

The traditional group followed the procedure written on the laboratory sheets to verify the known results.

The problem solving group utilized laboratory sheets that provided a series of possible solutions. This laboratory enabled the students to practice problem solving behaviors such as defining problems, forming hypotheses,designing experiments, interpreting data, and making conclusions or generalizations.

Results

Stepwise multiple regression was used to analyze the data. Analysis showed that three predictor variables together (logical thinking ability, treatment, and the interaction between them) accounted for 87.27% of the variance in the chemistry achievement. This is statistically significant at 0.05 level ($F=310$, $p<0.05$). Logical

thinking ability ($F=376, p<0.05$) and treatment ($F=9.65$, $p<0.05$), each made a significant contribution to the variation in chemistry achievement. Logical thinking ability accounted for 74.61% of the variance in chemistry achievement and treatment for 12.41% of the variance. But theinteraction between logical thinking ability and treatment did not account for significant portion in variation ($F=2.57$, $p>0.05$).

Logical thinking ability and treatment together accounted for 86.28% of the variance in science process skill achievement. This is statistically significant ($F=430.83$, $p<0.05$). Logical thinking ability ($F=735.68, p<0.05$) and treatment ($F=107.33$, $p<0.05$) each made a significant contribution to the variation in science process skills. Logical thinking ability accounted for 75.53% of the variance in science process skills, and the treatment for 10.74% of the variance. The interaction between logical thinking ability and treatment did not make a significant contribution to the variation in science process skills.

Discussion

Results showed that logical thinking ability is a significant predictor for chemistry achievement and science process skills because it requires formal operational reasoning. Abstract concepts in chemistry require students to use thinking abilities. So it is clear that these students at formal level can perform better in thinking processes.

Results indicated that the treatment accounted for significant portion of the variation in chemistry achievement and science process skills. The problem solving group had higher mean scores on the chemistry achievement test and science process skill test than the traditional group. Some chemical concepts require higher level Taxonomy. Problem solving approach is an underlying factor for development of these concepts because it enables students to perform problem solving activities such as identifying problems, formulating hypotheses, designing experiments, and making valid conclusion and generalizations.

References

Bender, D., & Milakofsky, L. (1982). College chemistry and Piaget: The relationship of aptitude and achievement measures. Journal of Research in Science Education, 19,3,205–216.

Chandran, S. ,Treagust, D.F., & Tobin K. (1987). The role of cognitivefactors in chemistry achievement. Journal of Research in Science Teaching,24,2,145–160.

Herron, J.D. (1975). Piaget for chemists: Explaining what good students can not understand. Journal of Chemical Education, 52,3,146–150.

Okey, J.R., Wise, K.C., & Burns, J.C. (1982). Integrated Process Skills Test–2. Department of Science Education, Athens.

Tobin, K., &Capie, W. (1981). Test of Logical Thinking. Department of Science Education,Athens.

Tobin, K., & Capie, W. (1982). Relationship between formal reasoning ability, locus of control, academic engagement and integrated process skill achievement, Journal of Research in Science Teaching,19,2,113–121.

PROMOTING CONCEPTUAL DEVELOPMENT IN YOUNG CHILDREN: THE IMPLICATIONS OF THE PRIMARY SPACE PROJECT ON CHILDREN'S LEARNING IN SCIENCE

T. Russell, & L. Mc Guigan, Centre for Research in Primary Science and Technology, University of Liverpool, Liverpool, United Kingdom

The Science Processes and Concepts Exploration (SPACE) Project is being conducted at two centres, the Centre for Research into Primary Science and Technology, Department of Education, University of Liverpool and the Centre for Educational Studies, King's College, London. The Project is based on constructivist principles that children's learning in science should take account of children's prior ideas. The research aims to identify the ideas that primary school children hold in a wide range of science concept areas and establish the viability of the techniques developed for eliciting ideas and managing subsequent interventions within classrooms.

This paper will draw on the research into children's understanding of Materials to illustrate the range of ideas children hold and identify the strategies teachers have used to develop these ideas.

The project has been run collaboratively between the University research teams, local education authorities and schools, with the participating teachers playing an active role in the Project work. The Project team aimed to have as much contact as possible with the teachers throughout the work to facilitate the provision of both training and support. Teachers were encouraged to incorporate the Project work into their existing classroom organisation so that they and their children were as much at ease with the work as with any other classroom experience. A large component of the Project work was classroom-based, and all of the children in the participating classes were involved as far as possible. Small groups of children and individuals were selected for additional activities and interviews to facilitate more detailed discussion of their thinking.

The research work was organised into two major phases each followed by the collection of structured data about children's ideas. These consisted of an elicitation phase designed to establish the ideas children held about a particular concept area and an intervention phase which explored the extent to which it was possible to encourage a change in the ideas in a direction which will help children develop a more scientific understanding of the topic. Using individual interviews of a sample of children complementary sets of data were collected prior to and

following the intervention phase. The data were analysed to identify changes in ideas across the sample as a whole and within individual children.

These phases of the Project form a coherent package which provide children with opportunities to explore and develop their scientific understanding as a part of classroom activity, and enables researchers to come nearer to establishing what conceptual development it is possible to encourage within the classroom and the most effective strategies for its encouragement.

In its final phase, the Project moved on to develop curriculum materials based on the classroom research. The issues which arise in moving from research to the development of curriculum materials will be discussed.

MAP SKILLS, A BASIC TOOL?

H. van Dijk, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Introduction

In 1991 a new research program was started at the Amsterdam Free University to investigate what specific subject differences there are between children and how education can be organized to reduce these differences. This research program is called "Curriculum studies and methodological differentiation in secondary education". The research is directed towards two main questions:
A.What differences in subject knowledge and skills do secondary school students show when they perform subject specific tasks?
B.Which methodological approaches help to improve the quality of education in the sense that more students achieve the goals formulated beforehand?

Geograhpical knowledge

The general objective of Dutch geography teaching for students between 12 and 15 years of age is: "Geography teaching is designed to equip young people with skills which will enable them to participate in various social and life situations in the spatial relationships of which they are a part: their own environment, the Dutch, the European and the world communities" (Commissie Herziening Eindtermen 1991, p.54). In other words, students need to acquire skills that will help them perform in a wide variety of regional frameworks. The research presented here investigates the differences that students manifest in answering questions for which such a set of geographic instruments is needed. In our research subject skills are defined as geographical procedural knowledge. Procedural knowledge in geography concerns either the actions which are necessary to pass from geographical facts to geographical generalizations, or those necessary to explain or predict facts using the geographical generalizations offered. Connected generalizations form conceptual knowledge (Van Westrhenen 1989a). Knoers (1973) states that education will prove more efficient when in addition to factual knowledge there is also conceptual knowledge. Conceptual knowledge offers possibilities for more and various kinds of achievement, because this kind of knowledge is more flexible than factual knowledge (Van Parreren 1974).
As far as learning processes are concerned this research is connected with the cognitive approach to learning processes as it is described by, among others, Boekaerts (1987) and Boonman & Kok (1982).

Geography teaching presents knowledge in various ways: texts, tables, graphs and maps. For the geographer the map is the most important of these various sources of information. Maps show lots of data and in this way offer the possibility of both finding locations, spatial distributions and areal differentiations and the areal associations between these spatial distributions and areal differentiations, which is important in order to achieve more insight into reality. What a user sees on a map and can conclude from that depends on his ability to use map skills. Map skills can be divided into three stages of action:

- Map reading: dentifying and naming phenomena on a map.
- Map analysis: recognition of spatial distributions, areal differentiation and areal associations on a map.
- Map interpretation: making statements or predictions using the areal associations discovered on the map.

Someone with limited map skills will generally only be able to describe the individual objects on the map. Greater map skills are necessary to organize information on the map, which makes it possible to find that relationships can be discovered between various characteristics of areas.
To answer the question about the differences in map skills between children a systematic description of the different kind of map skills is necessary.
However there is another reason why this description is necessary. In order to enable as many children as possible to achieve geographical objectives, it is desirable that they have an optimal motivation for their tasks. Motivation researchers such as Heckhausen and deCharms (see: de Bruyn 1979) point out that in order to get optimal motivation children should be able to set realistic goals and that the tasks should have a challeging character
(Heckhausen 1974). The level of challenge is mainly determined by its complexity. So complexity should be defined for every specific subject.

Research method

In this research complexity is defined in three ways.

- According to the character of the procedure (identification, classification and relating the several phenomena).
- According to geographical complexity. (the number of spatial distributions and area characteristics shown on the map).
- According to the character of the phenomena presented (fictitious or real).

Map 1 shows the most complex map with a difficult task. 45 children of about 13 year made the first maptest.

Results

Identification was not a problem for most children (90.40% correct answers). Correct classifications were made by one half of the group (50.88%). In making correct relations by finding the correct generalizations on the map 24.50% of the children were successful. Only few children were able to describe these generalizations.
Table 1 shows that the geographical complexity of the map influenced the results as well. Children are more successful in answering questions about simple maps than about maps where more phenomena are presented.

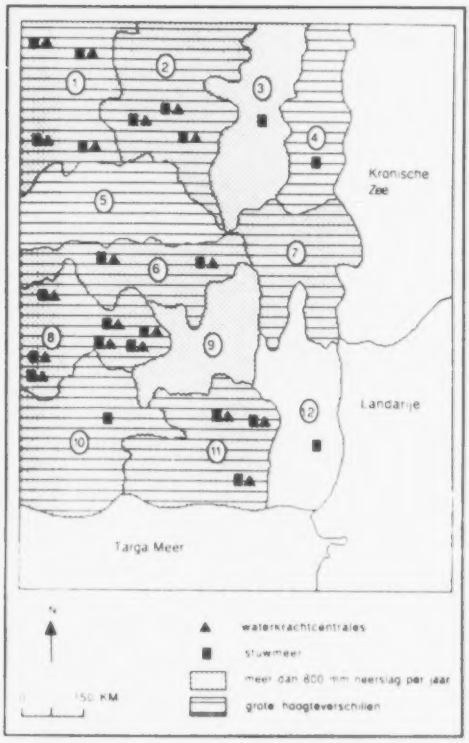


Fig. 1
Where would you built a new hydro-electric powerplant?

Table 1
Percentage correct answers related to the kind of mapskill and geographical complexity.

	Map relations	Map classifications	Map identifications
2 distr.+ 2 area diff.	14.25	33.00	80.50
2 distr.+ 1 area diff.	11.25	29.25	91.25
1 distr.+ 2 area diff.	24.25	56.00	78.75
2 distributions	28.00	62.50	91.50
1 distr.+ 1 area diff.	21.50	62.25	94.75
2 area differentiations	47.75	43.74	93.25
1 distribution		75.00	93.25
1 area differentiation		42.25	98.50

The effect of the character of the presented phenomena (fictitious or real) is limited. The average score on maps with real world phenomena is slightly higher than the score on maps using fictitious pheneomena.
When looking at the scores of the children related to the different kinds of maps, it indicates that the chosen definition of complexity of maps is successful. This means that these kinds of maps form a successful measuring-instrument for differences in map skills between children. It also means that this kind of tasks can be used in a system where children can set their one goals and choose a task with an optimal complexitylevel.

References

Bergen, Th. C. M. (1979). De veranderbaarheid van taakgebonden motieven in onderwijsleersituaties. In: Bruyn, E. E. J. de (red.) Ontwikkelingen in het onderzoek naar prestatiemotivatie. Lisse: Swets & Zeitlinger B.V.

Boekaerts, M. (1987). Psychologie van de leerling en het leerproces. Nijmegen: Dekker & van de Vegt.

Boonman, J.H., & W.A.M. Kok (1982). Het verwerven van kennis. Losbladig onderwijskundig lexicon, PO 4240, pp. 1-18.

Commissie herziening Eindtermen (1991). Nader advies kerndoelen voor de basisvorming in basisonderwijs en voortgezet onderwijs. Den Haag: Ministerie van O&W.

Heckhausen, H. (1974). Lehrer-Schuler-Interaktion. In: F.E. Weinert, C.F. Grauman, H. Heckhausen & M Hofer (red). Funk-Kolleg Peadagogische Psychologie I. Frankfurt am Main: Fischer.

Knoers, A.M.P. (1973). Leren en ontwikkeling. leerpsychologie ten behoeve van het onderwijs. Assen: van Gorcum.

Parreren, C.F. van (1974). Het functioneren van leerresultaten. In: C.F. van Parreren, & J.Peeck. Informatie over leren en onderwijzen, pp. 114-131. Groningen: Wolters-Noordhoff.

Westrhenen, J. van (1989). Regionale kennisoverdracht in nieuw perspectief. In: G.A.Hoekveld, & G.Schoenmaker & J. van Westrhenen (red). Wijkende grenzen, pp. 45-70. Amsterdam: KNAG/Vrije Universiteit.

INFORMATION PROCESSING IN AQUIRING AND USING GENERALIZED KNOLEDGE

Bodo Krause, Humboldt-Universität zu Berlin, Germany

Introduction

Information processing in acquiring and using generalized knowledge is based on memory content (pre-knowledge) and learning. Recent findings in cognitive psychology (e.g. Klix, Gentner, Herrmann) establish the structure of memory content consisting of objects (concepts) and conceptual relations. A main distinction in literature regards the mode of relationship. There are the attribute-relations which link concepts with their features or attributes and the event-based relations in form of n-p]aced predicates which are connections between concepts based on an event-relatedness. Both modes differ with respect to its way of storing and activation. Attributes and features are stored within the concepts, attribute relations are activated by comparison processes and form into a hierarchical structure of concepts. Event-based relations are stored in its context as semantic structures which will be activated simultaneously.

Learning, especially cognitive learning, denotes a class of information processing which results in the acquisition of knowledge with potential relevance for behaviour. Thus the result of a learning process is a new or modified memory structure that allows subjects to act or react appropriate. This means that subjects acquire knowledge and use it as a basis for behaviour. A main class of this knowledge is the recognition of invariances which allows efficient behaviour (i.e. minimal effort in storing and recall, simple and efficient decision structures). Therefore we ask for the principles and modes of acquiring and using invariances in dependence on different contents of invariances. More precisely: If we differentiate between attribute-relations an~ event-based relatedness then there should also exist differences in the information processing in learning and transfering this knowledge. Moreover it should be possible to distinguish between different learning procedures for acquiring the different types of invariances.

Method

To analyse the information processing in acquiring and using invariances we choose a compcsite experimental design consisting of following main components:

- a pre-test for proving homogeneity of experimental groups,
- a learning phase where subjects learned invariances by means of different methods (learning by examples vs. learning of rules),

279

- a testphase where subjects should use the learned invariances for solving a problem task.

Starting from this experimental approach we used the following 2 * 2 design of experimental groups:

Learning method Typ of invariance	Learning of rules	Learning by examples
attribute relation	exper. group 1	exper. group 2
event based relation	exper. group 3	exper. group 4

This experimental design clarifies the hypotheses under consideration regarding the differences in information processing between the different learning methods and the type of invariance both for the acquisition and the use of invariances.

Results

The experimental findings show:

- Both types of invariances will be learned and promote advantages for solving the problem task.
- Under condition of rule learning the learning effect of attribute relations is higher than the effect of event-relatedness.
- Under condition of learning by examples event-based invariances will be acquired more efficiently than attribute-based invariances.
- Additionally all shown effects depend on the cognitive difficulty of invariance content.

Conclusion and discussion

The experimental findings show that acquisition and use of invariances depend on the content of invariance and of the cognitive difficulty. Thus, in agreement with Sternberg's demand to study learning processes in connection with transfer situations, it results that increasing cognitive difficulty originates higher cognitive effort in acquiring this invariances but also produces a higher effect in transferring it to problem solving. Moreover the acquisition depends on the mode of information offering where a correspondence with rule learning and event-relatedness and a correspondence between learning from examples and attribute relations is shown. Together with earlier results this is in agreement with theoretical propositions in cognitive psychology and can partly be explained by parallel information processing.

A STUDY OF SELF-EFFICACY FOR SCIENCE PROBLEM SOLVING IN A COOPERATIVE INSTRUCTIONALCONTEXT

Maria de Fátima Chorão C. Sanches, Department of Education, Faculty of Sciences, University of Lisbon, Lisbon, Portugal

Introduction

Social cognitive theory of action (Bandura, 1986), when applied to the field of education, offers a conceptual foundation to test the general assumption that students can be taught to become more self-regulated learners by enhancing academic self-efficacy. In this regard, the use of self-regulated instructional strategies is ascribed importance on the following double basis: the self-efficacy information they provide the learner with, on the one hand, and the enhancement of both a sense of control over the student´s behaviour (Zimmerman, 1989; Harris, 1990) and academic motivation, on the other.

Cooperative problem solving, while facilitating the sharing of information and resources (Johnson et al., 1975; 1984), provides specific opportunities for the students´ interaction and, consequently, additional self and other referent sources for efficacy comparison. Among self-regulatory instructional strategies, problem solving through reflective thinking appears as a fundamental skill in science education. The literature reviewed emphasizes the problem solving role in developping the students´ interest in and more positive attitudes toward science knowledge. The present study is also based on the convinction that skill in any of the complex operations of thinking does not have to be developed as a by-product of learning specific curricular contents nor independently of them. Indeed, teaching for the students´ intelectual development must start with the teachers´ intention to do so while planning instruction

Being part of a larger research project which purports to investigate the nature and components of the context that might be associated to the 5th graders´ motivation and achievement in science, this study focused on the role of self-efficacy regarding specific problem solving skills. De Bono´s (1975; 1976) cognitive model of problem solving was adapted in order to integrate the learning activities. Inserted in the natural sciences program for 5th graders, the topic of energy (concept, sources, transformations, and uses) was structured according to De Bono´s problem solving approach. Instruction occurred in a cooperative classroom arrangement.

Based on the previous framework, it was hypothesized that (a) students following instruction under this condition would perform better than students in regular

classroom instruction concerning problem solving activities that required hypothesis formulation, analysis, synthesis, and application to new energy problems situations; (b) self-efficacy for problem solving would increase and contribute to a differential performance in new problem solving situations.

Method

Subjects were students in four 5th grade natural science classes. Experimental (N=50) and control groups (N=49) were randomly assigned from four intact classes. Both groups were taught by the same teacher who accepted to collaborate and open her classes to the experimental program. A pre-post-test with a control group design was used to test the hypotheses. Classes were organized in small groups (five students) wich worked under a discovery and solving problem approach.

The experiment was conducted during the second term of the academic year and took about three months. For the experimental group, the study of energy was planned according to the cooperative problem solving approach. First, following a video presentation, the subjects were free to identify problems about energy. Later on, another set of problems developed by the researchers were presented to be solved by the experimental subjects along the unit. Each group worked independently of each other, and on their own. They were given feedback by the teacher only after reporting the solutions for each set of problems.

The control group was predominantly exposed to classes planned according to a didactic approach. However, experimental demonstrations, and visualization of slides on energy sources and transformations were included in addition to some small groups work during the last part of the unit. Indeed the control group was not entirely taught according to the traditional didactic method of instruction. This situation was imposed to by the teacher and accepted by the researcher on the grounds of professional ethic.

The measurement of self-efficacy and problem solving performance

Self-efficacy for problem solving was measured by presenting the subjects with six problems concerning energy and asking how sure they were about solving those problems correctly. The levels of perceived solving problem efficacy were rated on a.scale ranging from one to six. Perceptions of the problems dificulty were also measured through a six point scale ranging from very dificult to very easy. Academic performance was measured in the end of the learning unit through achievement tests containing problems on energy. The test was composed of a series of problems that were structured to require the students´use of a variety of problem solving cognitive strategies.

Results and Conclusions

The findings of the present study reveal that, in accord with the predictions, self-efficacy accounted for a substantial portion of variation in problem solving performance. ANOVA of the results showed a similar pattern in control and experimental groups. There were not significant differences in the pre test between the two groups. This similarity between control and experimental groups may be in part due to the fact that control group also performed some activies in a cooperative manner.

Significant gains were found regarding self efficacy and performance for both groups. However, differential increments favoring the experimental group were found. regarding the following: transfering learning gains to new energy problems resolution, hypothesis development, and analysis of the problem components. In addition, the experimental subjects also felt more satisfied with the learning science activities than their counterparts.

Interpreted in the light of Bandura´s self-efficacy theory, the results point out the relevance of self-efficacy for learning problem solving skills. Indeed, according to Bandura, self-efficacy regards not a general but specific capacities which can be developed through appropriate self-regulated learning activities. Practical implications pertain to teachers´ education. Indeed, there is need to enhance the teachers´ awareness concerning this domain of teaching, so that instruction be planned for developping the student´s multiple and differential academic self-efficacies.

References

Bandura, A. (1986). Social foundations of thought and action: a social cognitive theory, Englewood Cliffs, NJ: Prentice-Hall.

De Bono, E. (1976). Thinking action - teachers´ handbook - cort vi, Direct Services ltd.

Kahney, H. (1986). Problem solving.- a cognitive approach, Open University

Jackson, K.F. (1983). The art of solving problems, The Bulmershe-Comino Problem-solving project at Bulmershe College Reading, Berks.

Johnson, D.W., Johnson, R.T., Holubec, E.J., & Roy, P. (1984). Circles of learning. Cooperation in the classroom, ASCD.

Johnson, D.W., & Johnson, R.T. (1975). Learning together and alone Coopertion, competition, and individualization, Prentice Hall, Inc., Englewood Cliffs.

Zimmerman, B.J. (1989). A social cognitive view of self-regulated academic learning, *Journal of Educational Psychology*, 81, 3, 329-339.

MODELS OF INSTRUCTIONAL INTERACTION AND STUDENT'S INTERPRETATIONS

Sanna Järvelä, University of Joensuu, Finland

Introduction

During recent years, the teaching in schools has been severely criticized. The criticism of the presentday education emphasizes that the higher order cognitive skills such as learning and thinking strategies and problem solving skills are not adequately developed by the conventional forms of school teaching. As respond to the criticism new cognitive models of teacher-student interaction have been developed. Reciprocal teaching model has been used for developing reading comprehension. It is a strategy that promotes both comprehension of text and also comprehension monitoring. Positive results have been reached particularly in helping poor readers become more strategic (Palincsar and Brown 1984, 1986, 1988). The idea of reciprocal teaching has also used developing writing process (Scardamalia and Bereiter 1985). Cognitive apprenticeship model (Collins, Brown and Newman 1989) contains some interactive instructional methods promoting learning, such as modelling, coaching, scaffolding, fading and reflection. These methods have proved powerful in helping students to attain cognitive and metacognitive skills and to ensure a gradual shift from external guidance to independent application of strategies on the way to more mature learning. Collins et al. have developed a model for designing ideal learning environments involving four dimensions: content, teaching method, sequense of learning tasks and social context of learning. This model offers an appropriate framework for the development of powerful learning environments. The three main techniques are based on observation, guided and supported practice, and feedback. They aim at the acquisition of cognitive and metacognitive skills. Modelling involves the observation by the student of an expert who is performing a certain task; this allows the student to construct an appropriate mental model of the activities required for the skill. Coaching refers to giving hints and feedback with a view to improving performance. Scaffolding consists of providing direct support on the right current skill level while student is carrying out the task. Two other methods aim at making students explicitly aware of their own cognitive and metacognitive activities. Articulation refers to any technique that helps students to better understanding and make explicit their knowledge and problem-solving procedures. Reflection leads students to compare their own cognitive strategies and solutions with those of experts and other learners. Exploration intends to increase the learner's autonomy in skilled problem solving as well as identifying

and defining new problems. Collins, Brown and Newman present two principles relating to the sequence of learning tasks: progressive complexity and diversity and global before local skills.The following guidelines are important with a view to realizing a favourable social context for learning. Situated learning means that students should be given tasks and problems representing the diversity of situations to wich they have to apply their knowledge and skills afterwards. Organizing opportunities for observe and contact with experts, provides learners with readily available models of expert-in-use and expert culture. That helps learners built and refine a conseptual model of the task they are trying to carry out. Enhancing intrinsic motivation for learning as well as fostering cooperative learning through small group problem solving are also important.
Despite some very promising results from experiments used cognitive apprenticeship style learning (see e.g. Palincsar and Brown 1984, 1986, 1988; Scardamalia and Bereiter 1985; Schoenfeld 1985) there are some weaknesses that are not much considered. The scaffolding interaction is very intensive and good concentration is needed. Isn't that sometimes very tiring and even motivation decreasing for pupils? Does it need certain characteristics to benefit of it? When the learning interaction is analysed, it is important to clarify what are the conditions for the reciprocal understanding or what are the real features for the beneficial cooperation. Modelling is powerful method to reveal expert's internal thinking and problem solving strategies. Seldom is however considered the reciprocity of teacher's modelling and student's cognitive understanding or ability to benefit of it. Is it easy for every student to benefit teacher's modelling or are there differences among students? Traditional apprenticeship is very contextualized, because the skill is learned in certain situation and for the certain purpose. Consequently the joint work between expert and apprentice is always concrete and meaningful. When we think cognitive apprenticeship and problem solving, there exists no concrete reference for teacher and student, which could help understanding during scaffolding process. Therefore it must be asked do we have to concretisize the thinking process before the idea works? Is it generally possible to use apprenticeship model for the abstract domain learning? It is also very important to clear whether it is possible to develop cognitive apprenticeship on a way that the development of creative thinking is possible. It should be possible to think new alternatives and solutions for old tasks' pupils had once solved. That is not typical in the traditional apprenticeship, where they only learn one high developed, but very standard work process.
The common aim at the new cognitive models of teacher-student interaction is both to improve pupils' domain knowledge and also teach them to develop their own learning skills. The purpose is to provide students with a new way; with cognitive possibility to understand better. Anyway before these cognitive learning interaction models are able to apply in ordinary classroom, more experiments and studies must be done to analyse the theoretical idea. In any case these new cognitive models of learning interaction offers a starting point for every teacher to discuss new alternatives for developing more effective learning interaction.

Experimental model of teacher-student interaction
The empirical experimentation is an effort to analyse the theoretical model of cognitive apprenticeship. Also the teacher-student and student-student interaction

BEST COPY AVAILABLE

are analysed according to the reciprocal sosioemotional interpretations. On this experiment the idea of cognitive apprenticeship is applied to computer assisted learning. Logotclego environment was used for mediating modern technological thinking and problem solving for secondary school students. Students constructed and programmed a physical model for the automated washing machine. Logo programming language used for programming and lego bricks (Technic Lego for computer control) for constructing the model. During the planning process of the program students used different extrinsic cognitive tools like tree diagrams in order to model the complex structure of the control problem. The experiment consisted of three parts of lessons kept in a computer classroom. One lesson was three hours. Students worked in pairs for programming a model for the automated washing machine. The whole project progressed as follows: The first lesson contained introduction, discussion about homework, construction of the model (washing machine) and exploration with the model. The second part included reproduction of previous lesson's topic, discussions about homework, introduction to using tree structure for problem solving and programming the model. The last part of lessons contained again reproduction of previous lesson's topic and programming the model in a new way. The learning interaction based to the principles of cognitive apprenticeship. The teacher used modelling and scaffolding as didactic principles, also the idea of reflection and articulation was used. Two kinds of teacher's role arranged: during one period the teacher was expert-teacher, and during an another period the teachers were expert-teacher and tutors, who already had some knowledge about legologo.

Method and results

The subjects of this experiment were all together 44 students from the seventh class of a Finnish secondary school. Their average age was thirteen years. A more detailed qualitative analysis comprised eight students. Before starting the experiment students filled in a questionnaire that measured their motivational and emotional interpretations during traditional classroom interaction. After the experiment they filled the same questionnaire to clear what kind of emotional and motivational interpretations they had after the new model of learning interaction. This material will used especially when analysing student's socioemotional interpretations concerning the learning interaction. One challenge for this experiment is to develop qualitative on-line methodology, which makes able to explicitly analyse teacher's and student's cognitive and socioemotional interpretations during the learning interaction. In order to analyse the quality of the learning interaction, three lessons in the two classes videotaped on a certain way. The special video tecnology was used, where the display card changed the computer screen to video signals. This made possibly to get simultaneously videotape of the interaction and also a real time videotape of the screen. All the elements of situation, videotaped interaction, voice and real time videotape of the screen, combined to the one complete tape. The data was collected by videotaping four student pairs working so, that it was possible to follow the interaction between students and the interaction between the teacher and students. Also the whole classroom situation with all pairs was videotaped . After collecting the video data, some pairs were interviewed. The interview was made by using a stimulated recall method, letting students to watch the tape where they

themselves work for programming. They were asked to tell spontaneously, while watching the tape, their feelings and thoughts about interaction. Some questions presented concerning for example teacher's modelling and scaffolding. The special tape, where all the elements of the interaction were connected, made possible for students' to report their feelings of the interaction. In addition, they were able to see their actual cognitive work concretisized on a computer screen and so reveal their mental processing reflected to external interaction. The results describing student-student and teacher-student interaction will be presented in the paper. With the help of explicit analyses of the meaningful reciprocal interpretations cognitive apprenticeship method functioning will be demonstrated.

References

Collins, A., Brown, J.S., & Newman S. (1989). Cognitive apprenticeship: teaching the craft of reading, writing, and mathematics. In L. Resnick (Ed.) Knowing, Learning, and Instruction: Essays in honor of Robert Glaser. Hillsdale, NJ: Laurence Erlbaum Associates Inc.

Palincsar, A.S., & Brown, A.L. (1984). The reciprocal teaching of comprehension monitoring activities. Cognition and Instruction, 1, 117-175.

Palincsar, A.S., & Brown, A.L. (1986). Interactive teaching to promote independent learning from text. The Reading Teacher, 4, 771-777.

Palincsar, A.S., & Brown, A.L. (1988). Teaching and practicing thinking skills to promote comprehension in the context of group problem solving. Remedial and Special Education, 9, 53-59.

Scardamalia, M., & Bereiter, C. (1985). Fostering the development of self-regulation in children's knowledge processing. In S.F. Chipman, J.W. Segal & R. Glaser (Eds.) Thinking and learning skills: research and open questions. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schoenfeld, A.H. (1985). Mathematical Problem Solving. New York: Academic Press.

COMENIUS' "ORBIS PICTUS" AND SOME EXPERIMENTS IN THE CURRENT USE OF PICTURES IN FOREIGN LANGUAGE INSTRUCTION

Iva Pychová, Silesian University, Karviná, Czechoslovakia

Introduction

Every country prides itself on its genius: England loves Shakespeare, Czechoslovakia pays tribute to Jan Amos Comenius, whose 400th anniversary of birth was commemorated in March 1992. Interestingly, the works of both these men have outlived their times due to the same qualities - their shining humanism and their art: Shakespeare's dramatic art and Comenius' ars docendi, the art of teaching.

Comenius "Orbis Pictius"

Of all the outstanding books by Comenius "Orbis sensualium pictus" (1657) excels, first, by introducing pictures as aids in foreign language instruction, an aspect which is usually most valued. Secondly, Comenius deliberately employed pictures for developing the learners' personalities, which is a novelty that has rarely been recognized. Thirdly, the content of "Orbis Pictus" consists of the most advanced knowledge of the first half of the 17th century. At a time when textbooks contained extracts from books by classical writers, the pictures of practical objects of everyday life must have captivated students all over Europe. Moreover, the encyclopedic character of "Orbis Pictus" and its logical system of trinity, employing the magic number of 3, made "Orbis Pictus" a model of rational thinking of its time and a model for all textbooks using pictures (e.g. Basedow's Elementarwerk).

The experiment

Although ample use of pictures can be found in contemporary textbooks, the employment of pictures is sometimes less functional and the content, form and layout of the pictures often miss any psychological, methodological, graphic or even logical rules or principles. This situation was a stimulus to our experiment, the aim of which was to discover the attitudes of learners to various types of pictures and to their use in foreign language instruction.

The experiment was conducted during 1990 and involved 46 secondary school students (aged 14-18), 41 School of Business undergraduates (aged 19-20) and 45 adult in-service teachers (aged 30-50) attending an intensive English requalification course. The experiment consisted of 10 tests investigating the

learners attitudes to picture content, colours, form, style, mode, drawing technique, focus and order of observation.

The experiment was a part of foreign language instruction and the subjects were not aware of being tested. However, they were always given enough time to consider their choices and were asked to make the selection responsibly. To stress the importance of the task the learners were afterwards asked about the reasons for their choices. The reasons were later used for the interpretation of the results. To increase the reliability of the tests each item was tested twice (with different pictures and within a certain lapse of time) and the results were then analyzed using appropriate statistical methods. In the preparation of the tests we built on extensive theoretical and research literature, both Czech and foreign (Arnheim, 1960, Gauvenet, 1966, Juva, 1966, Kunovjánková, 1988, Macek, 1976, 1987, Slavin, 1971, Spáleny, 1960-61, Stevick, 1986, Uzdil, 1988, Vernon, 1966, Wright, 1976, Zankov, 1955 etc.).

The 1st test measured the preferences the learners expressed in the choice of either verbal or picture tasks. Table 1 showing the results reveals a distinct preference for picture exercises (83%) with the learners, but especially with students (91%) and adults (87%).

In the 2nd test the learners judged the variety of colours in the pictures. Although there were slight differences between students (33% for multi-coloured pictures) and adults (40% for black and white pictures) the greatest number of all learners (48%) preferred four-coloured pictures. The choice might have been influenced by the content of the pictures, even though they did not differ very much.

The third test measured the choice between realistic drawings and a pictograph sequence. While students' answers did not express much difference, some undergraduates considered pictographs childish and naive and preferred realistic pictures. With adults the results were more clear cut: 78% in favour of pictographs. Total results showed 58% of preference for pictographs.

The 4th test assessing the effect of content of pictures (people or things) resulted in almost unanimous preference for pictures of people (73.5%).

In the 5th test the comprehensive versus "meagre" content of pictures was tested. Almost 70% of learners preferred less comprehensive, "meagre" pictures, especially adults (80%).

In the 6th test we intended to find the difference in the effect of marked (e.g. humorous) and unmarked (neutral) pictures. As it was difficult to find an unmarked picture (the picture form itself is a kind of markedness), we tested the effect of higher and lower marked pictures instead. Even if higher marked pictures were much preferred by the learners (92%), there was certainly a limit beyond which the positive effect of high marked pictures changed into a negative one. This limit remains to be determined.

The 7th test examined the effect of the drawing technique of pictographs, i.e. thick outlines and thin outlines. This time two versions of the same pictures were used and the choice to be made by the learners was much more detailed. The results showing the preference for the thick outline pictures (67%) were reached mostly owing to adults (84%), while undergraduates' opinions were divided (49% for thick outlines, 51% for thin outlines). This fact might have been influenced by the technical education of the undergraduates which made them prefer thin, technical looking outlines of pictures.

This technical bias seems to be supported by the results of the 8th test, which tried to find the learners' judgement of incorrect ratios of objects in the pictures. Although 94% of the learners did not identify the ratios as incorrect and illogical, 12% of undergraduates marked this distinction.
The 9th and 10th tests examined the focus and the order of observation of pictures. The results of the 9th test show that in the tasks administered, 36% of learners focused their attention on the picture in the middle, 39% on the top left picture, and 23% on the top right picture. Of all secondary school students 50% preferred the top left picture, probably because of its marked content (a picture of Vinnetou).
The 10th test included 8 items distributed throughout the picture. For the purpose of analysis, the order of observation was divided into four possible ways: horizontal, circular, vertical and random. The greatest number of the learners (45.5%) chose the horizontal order starting top left. While 27% of the learners chose a circular and 20.5% a random orders, a vertical order was chosen by only 7%.

Results

The results show a great preference for pictures compared with verbal exercises, especially for 4-coloured pictographs showing people in some marked situations. The pictures should be "meagre", drawn with thick outlines, the rations of objects need not be correct. Items noticed most lie in the middle or top left of the picture, the order of observation of the picture is in keeping with the European order of reading and writing (i.e. horizontal, left-right).
The most clear-cut attitude towards pictures is that of adults, which is identical with the general characteristics described above. The attitude of undergraduages is almost identical with that of adults, however, they also incorporate verbal exercises, pictures of objects, unmarked situations, thin outlines. Some of them marked incorrect ratios as illogical. The least clear-out is the attitude of secondary school students who liked both 4- and multi-coloured pictures, realistic pictures and pictographs, "meagre" and comprehensive pictures, thick and thin outlines of pictures. Their order of observation is both horizontal and circular. This attitude might stem not only from the immaturity of the students' opinions, but also from their greater need for pictures of all types. On the other hand, the attitude of adults might express a developed stage of need for visual representation of reality. These pictures should be coloured but easy to perceive (not distracting), reduced to a few essential items (easy to understand) and functional (easy to remember). The pictures should be drawn with thick outlines (easy to imagine) and their content should concern people in marked situations (easy to identify with). To learn the message of picture one should read it horizontally, left to right, starting top left.

Discussion

The reason for a preference for picture tasks compared with verbal exercises can be seen in a certain pre-processing of reality in the pictures, which facilitates and accelerates cognition, comprehension and learning. The transformation of reality in the picture (e.g. its simplification, separation of essential elements) supports retention and recall. Also simultaneous and overall perception of pictures as well as their provision of "ready-made", concrete images of notions makes cognition easier and requires less effort from the viewer. Consequently, the viewer's

concentration of attention can be longer and the intensity of his perception is higher (Pýchová, 1991). Four-coloured pictures are most preferred probably because their colours do not distract the viewer's attention as was proved with multi-coloured pictures (Allan, 1985). "Meagre" pictures make perception and cognition easier both for the so called "field-dependent" and "field-independent" viewers. In pictures with a limited number of items the viewer can easily "break free from the patterns and structures" of the picture (Dickinson, 1987). The preference for pictures of people may be attributed to the possibility of a viewer identifying him/herself with the persons in the picture. Objects in the pictures are usually judged by their functions in the depicted situations, therefore their size may be less relevant. It has been proved that thick outlines of pictures provide a feeling of simplicity and security. The perception of these pictures is easier than with thin outline pictures, which, however, give a feeling of precision and subtlety (Arnheim, 1960, Uædil, 1988). Focusing the viewer's attention on the middle of the picture has been explained in many ways. Research in visual perception (Vernon et al., 1966) supports the idea of the viewer's search for security in symmetry and in the "golden mean". Marked pictures stand out from among other pictures because of their unique, i.e. exaggerated, humorous etc. content or form. These attract more attention, arouse a greater interest and stronger emotion from the viewer than neutral pictures.

Conclusion

The results of the tests referred to above show definite preferences for pictures of certain qualities. Therefore they stress the necessity for teachers, textbook illustrators and publishers to take heed of these preferences, and to pay greater attention to the selection of pictures for foreign language instruction and education. After 400 years one must admire the intuition of Comenius in using this kind of pictures.

References

Allan, M. (1985). Teaching English with Video. Longman.

Arnheim, R. (1960). Art and Visual Perception.Berkeley. California.

Comenius, J.A. (1970). Orbis Pictus.Opera Didactica Omnia. Amsterdam, 1657. In: Dílo J.A.Komenského, Praha, Academia.

Dickinson, L. (1987). Self-Instruction in Language Learning. Cambridge, Cambridge University Press.

Gauvenet, H. et al. (1966). Le francais par la méthode audiovisuelle. Didier.

Juva, V. (1966). Pedagogický princip názornosti. Brno, UJEP.

Kunovjánková, D. (1988). K pojetí zásady názornosti. Pedagogika, 38, No.1, pp.7-25.

Macek, Z. (1976). K problematice obrazu pro jazykové vyucování a moznostem objektivizace jeho tvorby. In: Glottodidactica Olomucensis,I,Praha.

Macek, Z. (1987). Vyzkum pedagogické úcinnosti vyukovych videoprogramu, Pedagogika, 37, No.6, pp.659-665.

Pýchová, I. (1991). K funkci vizuálií v rozvoji osobnosti záka. Pedagogika, 40, No.6, pp. 669-684.

Slavin, A.V. (1971). Nagljadnyj obraz v strukture poznanija. IPL Moskva.

Spáleny, E. (1960-61). Funkce obrazu ve vyucování cizím jazykum. CJS. 4. No.3, pp. 105-113.
Stevick, E.W. (1986). Images and Options in the Language Classroom. Oxford University Press.
Uzdil, J. (1988). Mezi umením a vychovou. Praha, SPN.
Vernon, M.D. et al., (1966). Experiments in Visual Perception. Harmondsworth, Penguin Books.
Wright, A. (1976). Visual Materials for the Language Teacher. London, Longman.
Zankov, L.V. (1955). Opyt issledovanija vzaimodejstvij slova i nagljadnosti v obucenii. Sovetskaja pedagogika. Moskva.

IS MEANINGFUL ALWAYS FRUITFUL? THE SENSE OR NONSENSE OF EXERCISES

P.J. Meijer, & A.G. Sciarone, Delft University of Technology, Applied Linguistics Unit, Delft, The Netherlands

Exercises, meaning aids to learning are, and will remain, an important part of T2/VT curricula. Closed exercises such as gap–filling exercises are in current use as well as open exercises that accompany assignments in talking and writing. The general attitude about what constitutes an effective exercise–form has, at least by linguistics experts, undergone a number of changes over the years. Mechanically drilling by rote, intended to fix grammatical structures of the study–language in the mind by repetition, finally gave way to meaningful and contextualized exercises which approximate real communication. Exercises should consist of coherent texts in which a good understanding of the meaning of the sentence, and its context, is necessary so the assignment can be done. Ultimately, exercises must be offered to stimulate real communication (cf.Walz, 1989). In the opinion of some people it is exclusively the latter type which is effective. Attention, they feel, should not be on the linguistic structures primarily, but on the execution of pedagogical tasks with the help of the study–language. Attention to form, in their eyes, while important, is only secondarily so and then in the form of feedback or some other reaction.
The question of how a wide–scope exercise form – namely, the fill–in exercise – contributes to gaining fluency in a study–language was sought in a series of experiments. The way in which such exercises are composed has to satisfy the demands currently being put upon this kind of closed exercise. For this purpose, use was made of a CALL–system (Computer Assisted Language Learning) programme that consists of the following parts:

- an automatic homework checking programme that tests whether one of the requirements for such exercises, the learning of homework, was satisfied, and,
- a practice programme.

Task b) consisted of filling in missing words. These exercises differ from other similar exercises in that the student is confronted with a diversity of 'problems' instead of a single grammatical problem. During the classroom lessons, attention was given only to grammar, when grammatical aspects were hampering a good understanding of the texts. The material used in the lessons could be characterised as a natural method, in which grammar, right from the beginning, is made an integral part of the course: implicitly in the text, explicitly in a mini–grammar used as a reference grammar. Practice during the lessons, for the main part, consisted of free conversation on the basis of the material studied: forty–five

coherent texts in total from which students learned the 1200 most frequently used words in the Dutch language.
In Sciarone & Meijer (this congress) the average results are given relating to several of these experiments. This paper gives the results of a closer analysis of observations on several experimental groups (for a more detailed report see Meijer, 1990). The subjects of the experiments, who varied in age from 18 to 25, were students of various mother tongues who, judging from their evaluation tests, had no, or hardly any, knowledge of the Dutch language. They were put at random into different course groups and during the next three weeks followed an intensive course in Dutch. All the groups followed the same programme except the experimental groups who, during the course, also had special practice tasks. The last–named group were told that their results from the practice programme would influence the selection of their subsequent course. After the course was finished the general language fluency was assessed in a cloze test consisting of 100 items. In Table I, results are given of two experimental and two control groups. These two experimental groups did a series of gap–filling exercises of a more *general* character, in contrast to later groups: a variety of lexicological and grammatical problems were practiced. Differences in the results in the general language fluency test between the groups appeared to be insignificant.

Table 1
Group Scores for Evaluation Test and
General Fluency Test, and Standard Deviations
($n_{exp} = 26$; $n_{contr} = 30$)

	Evaluation Test		General Fluency Test	
	X	sd	X	sd
Experimental Groups	13.5	15.79	66.92	14.91
Control Groups	13.62	17.89	59.0	17.11

Subsequently, the scores of students receiving below–standard averages in the daily examinations have been excluded from the end–result comparison. No significant differences in the average scores for these selected groups have been found. Neither making exercises nor better mastery of homework could explain why the experimental groups had only a minor, and not a significant, superiority on the final examination. Luck and ineffectiveness of exercises are rather difficult to accept and for that reason registrations of students performance in the practice program have been closely examined. These data, for each participating student and for every exercise, consist of the following:

- beginning and ending time of the exercise;
- the answers given by the students;
- the score attained by each student;
- the use of feedback.

These data may indicate that students involved in the exercise programme might have developed greater ability to do this kind of exercises in the course of the study. To find out, the data were analyzed according to certain indicators, both for groups and for individual students.
The findings were these:

- Group scores for items belonging to the same grammatical category showed a downward trend; this was true both for scores for first answers alone and for last answers (students were allowed to make several attempts to find a right answer);
- A minor positive correlation appeared to exist between the frequency of items offered in the exercises and the difference in group scores for first and last answers; in other words, offering items more frequently is hardly the way to make sure that the first answer will be the right answer. Note that students receiving below–standard average scores on daily testing have been excluded from both calculations.
- Scores of individual students showed wide fluctuations for series of grammatically similar items.
- For grammatically similar items, the number of wrong answers per item that preceded the last (and thus correct) answer, didn't decrease during the course; this counts for successful and unsuccessful students.
- To develop a coherent picture, ratios showing the number of correct answers to the number of items per exercise were calculated, and also ratios of the number of correct answers to the number of answers supplied by the student. The weak student requires more attempts before getting the correct answer. The difference between a weak student and a good student is smaller on the subject of the ratio of correct answers:items. The problem of low ratios seems to increase, rather than tapering off, the further the student progresses in the exercise, but increases more strongly for the weak student.
- The time required before finding the correct answer increases for weak students as well as for successful ones.

Students doing daily exercises did not do significantly better on a general language fluency test; neither they achieved better as they advanced in the series of tests.

In view of the apparent inability of students to judge the nature of their mistakes, demanding language analyzing abilities as a problem solving strategy seemed less reasonable than presenting an example, more or less conforming to what happens with TI acquirement. By means of a second experiment we wanted to observe whether a fill–in exercise of that character would be effective. In a new series of exercises, every incomplete sentence was preceded by an example sentence and subsequently the same exercise had to be done without the help of the examples. The items of these exercises contained only a restricted number of grammatical problems. The results in the general language fluency tests showed no significant differences between experimental and controlgroups. Group scores for the categories on which the exercises were concentrated showed a significant difference in only one category (see Table 2).

Table 2
Average Scores, Standard Deviations and Reliability Coefficients for Grammatical categories in General Fluency Test. ($n_1 = 16$, $n_2 = 20$)

		Group 1 (exp.)			Group 2 (contr.)		
		pbar	sdp	KR20	pbar	sdp	KR20
verbs	(k=30)	80.62	16.11	0.81	70.16	25.40	0.78
adjectivs	(k=7)	82.14	12.72	0.20	80.0	2.90	–0.36
negativs	(k=5)	75.0	18.75	0.64	57.0	16.43	0.65
conjunct.	(k=11)	80.68	14.10	0.47	75.90	15.13	0.39
"te"(+inf)	(k=4)	81.25	22.24	0.49	81.25	18.87	0.52
Prenom.adv	(k=5)	42.50	25.54	0.71	48.0	36.33	0.34

Conclusions

No indications were found that either type of fill–in exercise, in general, contributed to the development of fluency; or that exercises which concentrate on certain grammatical structures stimulate such fluency; or that regular exercising increases the ability to do that kind of exercise. That is why, in our view, it is difficult to advocate adopting these exercises in programmes for T2/VT education. Additional research should establish whether more meaningful exercise forms can be developed to be used in their stead.

References

Meijer, P. J. (1990). Gap–filling exercises in the first phase of acquiring NT2. An analysis of CALL registrations. (Dutch), TU Delft Internal Report, Applied Linguistics Unit.

Sciarone, A. G., & Meijer, P.J. Does practice make perfect. On the effect of exercises on second/foreign language acquisition. In preparation.

Walz, J. (1984). Context and contextualized language practice in foreign language teaching, Modern Language Journal ,73.

COMPUTER CONTROLLED READING INSTRUCTION: TRANSFER MECHANISM IN VISUAL WORD RECOGNITION

Egbert M.H. Assink, Göran Kattenberg, & Michel Schmitt,
Department of Psychology, University of Utrecht, Utrecht,
The Netherlands

Transfer Effects in Reading Instruction

Learning to read fluently is a central skill in literacy acquisition. People seldom realize that children who successfully completed their first course in reading and spelling truly accomplished an extraordinary achievement. Normally, we do not realize the full impact of this accomplishment, until we encounter children who fail to master the intricacies of the alphabetic code system. The question why some children fail to achieve fluent reading has intrigued both researchers and educators for a long time. This also holds for the related question why specially designed training programs for poor readers produce course-specific learning outcomes. Trained subject specifically improve their word identification skills with regard to the materials that were dealt with in the training and fail to generalize the practiced skills to analogous (e.g. Reitsma, 1988, Van Daal & Reitsma, 1989). Recently, positive results have been reported in two experiments using onset and rime units as transfer mechanisms in reading (Goswami, 1990). In the first experiment, children learned more about common consonant blends at the beginnings of words (*trim-trap*), which constitute the onset, than at the endings of words (*wink-tank*), which break up the rime. In contrast, in experiment 2 adding a common vowel led to much greater improvement in learning about spelling sequences at the ends of words (*wink-pink*), which now constitutes the rime, than at the beginnings of words (*trim-trip*), which extend the onset. A major conclusion of these experiments is that phonological categorisations do seem to play a role in learning about spelling sequences.

The aim of the present study was to establish if transfer effects in word identification skill could be obtained by training poor readers in their processing speed of intra word orthographic clusters. The task used in our experimental training was a semantic decision task using word pairs varying in orthographic complexity. We will briefly describe this task here. A more complete account may be found in Assink & Kattenberg, (1991). The semantic decision task used in our experiment was similar to the task described in Jarvella & Snodgrass (1974). Subjects were asked to compare the meanings of two nouns in Dutch. A pair always consisted of a singular and a plural form. The subject had to decide if the members of the pair had identical stem morphemes (roots) or not, (e.g. compare

book-books versus *brook-books* in English). Essential for the experiment were the various classes of singular--plural pairs, as defined by the degree of similarity, or congruence, in orthographic structure between both members of the singular--plural pair. In a previous reading level design study reported in Assink & Kattenberg, (1991) poor readers turned out have more difficulty in processing incongruent pairs, showing intra word orthographic dissimilarity.
We designed a computerized training program aimed at accelerating word identification skill focusing on orthographic processing at a subword level. Subjects were trained in classifying the various singular-plural types present in Dutch orthography. These classes are specified below. There were two instruction conditions. The first one focused on the identification of orthographic cues by presenting the crucial letter cluster alternations involved. The second training focused at speeding up classification by presenting abstract graphic symbols as category labels. Since both types of training were focused on providing a generalized strategy for the analysis of orthographic structure, transfer effects might be expected. We were interested in the effectiveness of both types of training, particularly regarding their difference in producing transfer.

Method

Subjects

S. were 30 poor readers attending a school for learning disabilities to improve their reading skills. Reading level of these Ss. was assessed with a (Dutch) standard reading achievement test (Brus & Voeten, 1972). Reading test scores were used to assign the Ss. to 3 matched treatment groups of 10 Ss. each. Two groups were trained in an orthographic classification task. The first group was trained to detect critical letter features, the second group learned to classify the same materials with the help of abstract graphic symbols. The third group got no training and served as a control. The Ss. had a mean age of 12.6 yrs and a mean reading age of 9.4 yrs.

Materials

The following types of singular-plural pairs were used as items:

- Items representing perfect sound-spelling correspondence, e.g. *boek-boeken* 'book-books';
- Items showing congruent orthographic structure, but phonologically incongruent in the singular, e.g. *paard-paarden* [pa:rt/pa:rd n] ('horse-horses');
- Phonologically congruent , but morphologically incongruent items e.g. *huis-huizen* 'house-houses';
- Similar to type 3, but with additional alternating vowel orthography, e.g. *baas-bazen* [ba:s/ba:z n] ('boss-bosses');
- Items showing reduplication of the final consonant in the plural, e.g. *fles-flessen* 'bottle-bottles';
- Items showing an alternating vowel orthography, e.g. *droom-dromen* [dro:m/dro:men] 'dream-dreams'.

Procedure and Design

Both instruction programs used were computer controlled. In the letter condition, pairs of nouns were presented, followed by a screen with orthographically

represented choice options. To illustrate, the pair *baas-bazen* was followed by the options:

- aas - aaz
- aas - aas
- aaz - aze
- aas - aze (correct option)
- aze - aas
- none (control item)

Ss. could use the mouse to click the correct option. In the graphic symbol condition, Ss. were first instructed to relate the various orthographic categories to specified symbols, e.g. a cross for type 1, a square for type 2, etc. In both instruction groups the training was preceded by a short teaching module about how to operate the mouse for making correct choices. In both conditions this pre-instruction also included an explanation of the choice options available for the student.

Ss. worked individually with the computer program in five daily sessions of about 45 minutes. Instruction was stopped when the student had reached a sufficient skill in categorizing items. The mastery criterion was 90% correct responses and a decision time of 3 seconds or less per item. All subjects in both instruction groups reached this criterion within a week. There were two computer controlled post tests: a course specific retention test and a transfer test. In the course specific test Ss. were presented a sample of items that had been used in the training sessions. The transfer test contained corresponding items that had not been used in the training ever before. In both tests Ss. were required to evaluate the semantic identity of the members of a pair. Decision times (accuracy of 1 ms) were registered by the computer program. In addition to these tests the standard reading achievement test that was used for group matching was administered once again.

Results

Separate MANOVAs were run for contrasting both instruction groups with the control group. Although graphics instruction group did perform very well and soon reached mastery during the training, their post test results were disappointing, both with regard to course specific and transfer learning results. The graphics group did not outperform the controls, neither on the course specific test $F(1,17)=0.02$, $p<0.89$, nor on the transfer test. In fact, this group even did worse on the transfer test $F(1,17)=9.83$, $p<0.006$.

In contrast, the letter group showed a main instruction effect, both on the course specific test $F(1,17)=4.59$ $p<0.04$, and also on the transfer test $F(1,37)=7.33$, $p<0.014$.Post hoc analyses of the main effect of instruction showed that this effect was mainly due to the morphologically regular types #1 (*boek-boeken*) and #2 (*paard-paarden*). The letter group also score significantly better on the standard reading achievement test post test.

Discussion

The most striking outcome of this study is that the two training groups strongly differed in their overall effectivity. The pattern of results obtained suggests that learning abstract linguistic distinctions like orthographic patterns might even be

obstructed by choosing an inadequate instruction mode. Our results suggest that learning about written word patterns may be accelerated and even produce transfer to new reading materials.
The ineffectiveness of the graphic symbols condition is similar to the results obtained by Bryant & Bradley (1983) in their study about phonological awareness and learning to read. In their experiment there was also an instruction condition working with formal conceptual classes not explicitly associated with orthography. This condition also turned out to be relatively ineffective. Our experiment shows that orthographic patterns may be used in reading instruction in order to obtain transfer in reading.
With regard to the superior performance of the letter group on the repeated standard reading achievement test we must be careful. Although the letter group was also superior on this test it is not fully clear if this must be attributed to our training. This result needs more detailed analysis, especially since the control group performed slightly poorer on this test, as compared to the obtained pretest scores.
Nevertheless, we think that our result are sufficiently interesting for further exploration. Especially the issue of transfer and instruction on sublexical orthographic patters seems a particularly intriguing topic for future research.

References

Assink, E.M.H., & G.P.A. Kattenberg (1991). The Use of Orthographic Structure in Normal and Poor Readers. Paper presented at the 1991 AERA Annual Meeting, Chicago, Il.

Bradley, L., & Bryant, P.E. (1983). Categorizing sounds and learning to read - a causal connection. Nature, 301, 419-421.

Brus, B.Th., & M.Voeten (1972). Eén minuut test [One-minute test]. Nijmegen: Berkhout.

Daal, V.H.P. van, & Reitsma, P. (1989). The Development of Orthographic Knowledge in Beginning Readers. Paper presented at the 6th European Conference on Reading, Berlin.

Goswami, U. (1990). Onset and Rime Units and Transfer in Reading. Paper presented at the AERA Annual Meeting, Boston, MA. (Also submitted to Child Development).

Jarvella, R.J., & Snodgrass, J.G. (1974). 'Seeing ring in rang and retain in retention, on recognizing stem morphemes in printed words'. Journal of Verbal Learning and Verbal Behavior, 13, 590-598.

Reitsma, P. (1989). Development of Orthographic Knowledge. In P. Reitsma, & L. Verhoeven (Eds.), Acquisition of Reading in Dutch. Dordrecht-Holland/Providence RI - USA: Foris Publications.

TEACHING STRATEGIES TO POOR READERS

Miriam Walraven, & Pieter Reitsma, Paedologisch Instituut Amsterdam, Duivendrecht, The Netherlands

Introduction

Instruction in reading comprehension has been a topic of increasing interest over the last few years. Many researchers have sought to study the effect of different kinds of training, varying from instruction in specific reading comprehension *skills* to instruction in comprehension *strategies*. (Reitsma & Walraven, 1991).

The present paper focus on the effectiveness of teaching reading comprehension strategies, also defined as metacognitive, monitoring or self-regulating strategies, which stimulate active and strategic reading, facilitate reading comprehension, and lead to self-regulation of reading activity (Palincsar & Brown, 1989). Although research on this topic has been relatively scarce until now, the findings have been very promising (Paris et al., 1984; Duffy et al., 1987).

However, little is known about the effects of instruction in strategies for special populations, such as reading disabled, emotionally disturbed children. It is widely acknowledged that a strong, perhaps intrinsic or reciprocal, relationship exists between reading problems and behavioural problems. In clinical practise, however, the latter are frequently considered of prime importance and are often attended to first. Reading problems and its remediation are consequently sometimes overlooked, although they are obviously of equal significance. Therefore the purpose of the present study is to determine the effectiveness of instruction in strategies for children with reading disabilities, who are admitted to a child psychiatric clinic.

The pupils were taught the following strategies for reading comprehension: clarifying the purpose of reading in order to determine the appropriate approach; activating background knowledge to create links between what is already known and the information in the text; making predictions about the text and evaluate them; using control activities by self-questioning in order to determine if comprehension is occurring; allocating attention to the major content and the main ideas; and summarizing and interpreting the information given in the text.

It was hypothesized that children who attain the experimental training program would increase their strategy knowledge, which in turn could have an impact on the process of reading comprehension, probably leading to better performances on reading comprehension tasks.

Method

Subjects

	control group	experimental group
age (years;months)	10;11	11;1
verbal IQ (WISC-R)	89.5	85.6
vocabulary (max.40)	35.0	35.9
decoding (words in 3 minutes)	209.1	209.3

Subjects were 24 children from two schools for special education associated with a clinic for intensive child care. They were selected according to the following criteria: age between 10-12 years (educational grades 4-6); scores on a standardized reading comprehension test lower than the mean score of a one-year younger norm group; minimal decoding skills at the level of grade 3; and minimal verbal IQ of 80.
Twelve pupils were selected as a control group, attending regular reading lessons, while twelve others received the experimental program. The groups were comparable for the variables as described above. No significant group differences were found on the pretests.
Although the design also provides two conditions in the experimental group, respectively strategy instruction including and excluding the component of activating background knowledge, in the present paper the results are aggregated over both conditions. The reason is that first the question has to be answered whether experimental training has any benefits for reading comprehension at all.

Instruments

An experimental questionnaire for reading comprehension strategies, describing 12 relevant and 3 irrelevant strategies which could be used before, during and after reading, was used to determine the level of declarative knowledge of these strategies (Maximum score 12). Data about the performance in reading comprehension were collected by the following tests: a standardized cloze-test, with maximum score of 50 and mean score of 34 for normal readers in grade 4, and an experimental test for reading comprehension, consisting of 4 passages of about 140 words with 8 or 9 true/false propositions (maximum score=35). All tests were administered both before the program started (pretest) and approximately one week after the intervention (posttest).

The program

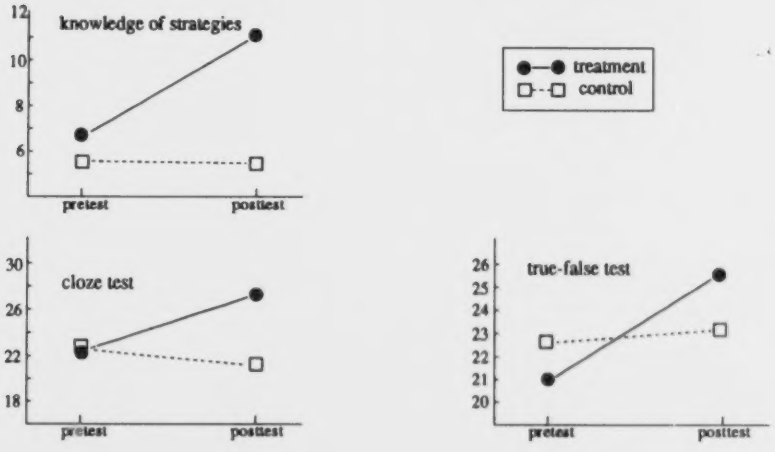
The experimental program consisted of 13-14 lessons, each lasting 30-40 minutes, which were attended by groups of 3 pupils. The instruction was based on principles of direct instruction and reciprocal teaching and was divided in the following phases: (1) repeating relevant knowledge from previous lessons; (2) explaining the aim of the new lesson; (3) modelling the use of a strategy; (4) guided practice; (5) independent practice; and (6) paraphrasing the new information of the lesson.
During the phase of independent practice the procedure of reciprocal teaching was followed, by encouraging pupils to assume the role of the teacher (Brown & Palincsar, 1989). In the reciprocal teaching procedure, children and teacher take turns leading the dialogue about the text and its meaning, breaking the text into

segments while applying the strategies. Responsibility for the comprehension process is shifting from the teacher towards the pupils, making them more active and self-regulating readers.
In the course of the program the following monitoring strategies were taught:

learning goal	pupil's strategy
determining purpose for reading	Why am I going to read?
activating background knowledge	What do I know already about the subject?
making predictions about the content	What is the text probably about?
flexible control of comprehension	Stop! Is this clear to me?
selecting important ideas	What is important to note?
summarizing and evaluation	What did I learn from this text?

The strategies were introduced step by step and were described in concrete language, visualised by an illustration. Pupil's strategies, accompanied with illustration, were presented on a "help-card"", each new card containing the strategies attended to before, completed with the strategy which is introduced and discussed during that phase of the course. The strategies were practised while reading expository texts, all age-appropriate and challenging.

Results
The results, summarized in the figures below, indicate significant interaction effects between repeated measures and group for all dependent variables.



Conclusion
The present study provides clear evidence for the possibility of improving reading comprehension by teaching strategies. Furthermore, the findings indicate that strategy training can be effective even for this special population of reading-disabled, emotionally disturbed children. This makes the present study of

particular interest for special education. More research in this area is required to investigate the possibility and effectiveness of implementation of the program in the actual practice of (remedial) teaching and to provide evidence about the relative effectiveness of the various components of the experimental program.

References

Brown, A.L., & Palincsar, A.S. (1989). Guided, coöperative learning and individual knowledge acquisition. In L.B. Resnick (Ed.), Knowing, learning and instruction. Essays in honor of Robert Glaser.Hillsdale New Jersey: Lawrence Erlbaum Associates, Publishers.

Duffy, G.G., Roehler, L.R., Sivan, E., Rackliffe, G., Book, C., Meloth, M., Vavrus, L., Wesselman, R., Putnam, J., & Bassari, D., (1987). Effects of explaining the reasoning associated with using reading strategies. Reading Research Quarterly, 22, 347-368.

Palincsar, A.S., & Brown, A.L. (1989). Instruction for self- regulated reading. In L.B. Resnick & L.E. Klopfer (Eds.). Toward the Thinking Curriculum: Current Cognitive Research. 1989 Yearbook of the Association for Supervision and Curriculum Development.

Paris, S.G., Cross , D.R., & Lipson, M.Y. (1984). Informed strategies for learning: A program to improve children's reading awareness and comprehension, Journal of Educational Psychology, 76, 6, 1239-1252.

Reitsma, P., & Walraven, A.M.A. (Eds.) (1991). Instructie in begrijpend lezen [instruction in reading comprehension]. RAIN 1. Delft: Eburon

COMPUTER-BASED READING AND SPELLING PRACTICE

Victor H.P. van Daal, Paedologisch Instituut, Amsterdam-Duivendrecht, The Netherlands

Reading-disabled children seem to have considerable difficulties in acquiring phonological recoding skills which are considered to be very important in becoming a proficient reader. It is hypothesized that providing speech feedback and integrating spelling activities within reading practice may support the development of phonological recoding skills. Several experiments were conducted in order to reveal which elements of computer-based reading practice were effective in this respect.

In experiment 1, the format of the speech feedback was systematically varied. For 17 subjects high quality digitized speech was available on call. When speech feedback was requested either whole-word sound was provided (n = 8) or segmented-word sound was presented (n = 9). A control group (n = 14) did not receive speech feedback at all. In a posttest both practised and nonpractised words were presented. Both help call behaviour during practice and changes in reading rate and accuracy from pretest to posttest were analyzed. It was found that speech feedback requests were not dependent on word difficulty. More calls for whole-word sound were made than for segmented-word sound, while latency times for whole-word sound requests were shorter than for segmented-word sound requests. Both forms of speech feedback were equally effective in reading practice words which were initially read incorrectly. However, there was a tendency for the learning effect on nonpractised initially hard-to-read words to be largest when segmented-word sound had been available. It is concluded that, whenever the goal of reading instruction is to memorize particular words, whole-word sound as well as segmented-word sound can be used. However, when more productive skills are aimed at, the most promising way of giving support is to present segmented-word sound, although only a nonsignificant tendency for transfer was found in the present study.

In experiment 2 the process of requesting speech feedback both in reading-disabled children and reading-age matched normal beginning readers was closely examined. Disabled readers and normal beginners were compared on requesting help in the form of speech feedback during computer-based word reading. It was also examined whether it is best to give feedback on all words or to allow the disabled readers to choose. Normal beginning readers and reading-age matched pupils with reading problems engaged in reading practice with speech feedback on call for both difficult and easy words. A set of both difficult and easy sums was completed as a control task. Another group of reading-disabled pupils

who were also matched on reading level practiced the reading of words with unsolicited speech feedback. In the first two groups the selectivity in requesting help in both the reading and arithmetic task was assessed. All subjects were tested on the reading of both practiced and nonpracticed words. The results indicate that both disabled readers and beginners were very sensitive to sum difficulty but not to word difficulty, though in the present study the students were more selective than in the previous study. The inclusion of more easy words might account for this finding. The beginners requested help only during the first sessions of reading practice, whereas the disabled readers remained dependent on the speech feedback. The reading-disabled pupils learned most when the computer unsolicitedly delivered the spoken form of all words. It is suggested that reading-disabled children do not have impairments in metacognition, but lack efficient decoding processes which are prerequisite for acquiring metacognitive skills. Therefore, remediation should first be focussed on decoding skills.

In experiment 3 a spelling task was added to the computer-based reading practice. To examine the effects of computer-based reading and spelling practice on the development of reading and spelling skills a pretest-training-posttest experiment was conducted in the Netherlands. Eleven girls and 17 boys with written language disorders (on the average 9 years, 7 months old and performing 2 grades below age expectancy) practiced hard-to-read words under three conditions: reading from the computer screen, copying from the screen, and writing from memory after presentation on the screen. For all words whole-word sound was available on call during practice. To assess learning effects both a dictation and a read-aloud task were administered in which nonpracticed control words were also presented. During training, the computer kept record of several aspects of the pupils' learning behavior. It was found that copying words from the screen resulted in significantly fewer spelling errors on the posttest than writing words from memory, and that both forms of spelling practice lead to fewer spelling errors than only reading words during practice. All three forms of practice improved to the same degree both the accuracy and fluency of reading the practiced words aloud. It is concluded that spelling should be subsumed in reading practice.

Finally, in a series of experiments the issue was addressed whether disabled readers can be trained to use within-word (multiletter) units in both recognizing printed words and in attacking novel written words. Several analogy methods were employed to instruct children with severe reading problems about multiletter units. In experiment 1 and 2 short lists of words that differ in only one letter were repeatedly practiced. In experiment 3 the subjects blended written and spoken words in which the place of segmentation was systematically manipulated. Subjects were tested in all experiments on the reading of practiced words, on nonpracticed words that both orthographically and phonologically resemble the practiced words, and on nonpracticed nonsimilar words. The findings indicate that (1) all types of practice were beneficial for the recognition of the practiced words, (2) transfer to the reading of novel words appeared only when graphemic and orthographic aspects of words were very intensively trained, (3) transfer effects were *not* based on improved handling of within-word units, but on a general improvement of decoding skill, and (4) at least in Dutch, segmentations according to the onset-rime principle are *not* more useful than segmenting written and spoken words at other boundaries.

AWARENESS OF TEXT STRUCTURE: A COMPARISON OF THREE METHODS TO ASSESS AWARENESS OF STRUCTURES OF EXPOSITORY TEXT

Ans Bergmans, Marinus J.M. Voeten, & Cor A.J. Aarnoutse, Department of Educational Sciences, University of Nijmegen, Nijmegen, The Netherlands

The first purpose of this study was to examine whether good and poor readers in the first form of secondary school are aware of four structures of expository text (collection, comparison/contrast, causation and problem/solution). The second purpose was to find out which of three measures of awareness gives the most information about individual differences. Partly, this study was designed as a replication of a study by Richgels et al. (1987). We used three measures of awareness, adapted from their study, namely a recall task, a composition task and an interview based on a matching task. The measure that has been most widely used, is written recall after reading a text. The basic assumption underlying this task is, that readers who are aware of text structures and who use this knowledge while reading, will reproduce in their recall the organisational pattern of the text they read. In this reproduction task memory will play an important role. The other two measures are supposed to make higher cognitive demands. The composition task is based on the idea that use of a particular structure to organise written compositions provides evidence for structure awareness. This production measure demands more initiative on the part of the students in applying their awareness of structure. The interview is supposed to make even higher cognitive demands because it requires metacognitive awareness of structure, or the ability to label text structures.

Our study was based on Meyer's approach of analysing text structures (Meyer, 1975; 1985). Her system describes organisational patterns found in text as hierarchical, tree-like structures. The highest level of the text is characterized by logical or rhetorical relations (the top-level structure of a text) that tie together large chunks of information. These relations signal particular text structures, like the four mentioned above. Meyer suggested that readers who are not aware of structure in text may employ a list-like way of encoding textual information. In contrast, skilled readers may use a structure strategy; they try to identify the top-level structure in order to recognize the gist of the text. This strategy facilitates comprehension and recall. Employing a structure strategy presupposes awareness of organisational patterns found in texts. Our expectation was that good readers are more aware of

text structures. We also expected that differences between good and poor readers are larger in tasks with higher cognitive demands.

Method

Subjects were selected from six classes of first-formers of a school of secondary education (Havo/Vwo). Twenty-five good and twenty-five poor readers were selected on the basis of a test for identifying the main idea of a text. Good readers were those who scored at least one standard deviation above the sample mean; poor readers were more than one standard deviation below the mean on the main idea test.

In the recall task we used two text structures namely causation and problem/solution. Each structure was represented in two texts consisting of four short paragraphs. In addition, of each text a scrambled version was made. The logical sequence of paragraphs was disturbed. Within each paragraph sentences were randomly ordered. In addition, words signalling the text structure were replaced. After reading the text the students had to write down what they remembered from the text. Every student was given four texts: a normal and a scrambled text for both structures. The order of presentation was counterbalanced.

In the composition task the same two text structures were used. For each of those a topic sentence was given and the first paragraph of a text to be completed. This first paragraph clearly indicated a causation or a problem/solution structure. All students had to write two compositions. Finally, in an interview the students had to explain their answers on a matching task which they carried out before. In this task they had to match texts with the same structure (causation, problem/solution, collection and comparison/contrast). In each of the four items of this task the students were presented with three passages: a target, a match and a foil. The three passages had different content. Each structure appeared once as a target and match, and once as a foil.

For the recall task four and for the matching task twelve texts were constructed. The difficulty level of the texts was regulated by controlling the length of the text, by computing a readability index and by controlling signaling of the text structure. Students' perception of their knowledge of the passage topics in the recall and composition task was assessed by a questionnaire.

Recall protocols and compositions were rated by judging to what extent the intended structure had been carried through by the student. The main points in scoring the interviews were whether the student succeeded in finding the correct match and whether the student could explain this with arguments pointing to the structure of the text.

Results and conclusions

The results showed that students from the first form have some awareness of text structures. Rather large differences in the expected direction were found between good and poor readers on all three tasks. There were no consistent differences between the four text structures. The results from the recall task showed that disturbance of the text structure negatively effects the reproduction. Contrary to our expectation, this effect was alike for good and poor readers. The results on the recall measure and on the composition task did not depend upon differences between students in topic familiarity.

Consistent with previous research the three measures gave different results. The intercorrelations between the three measures were rather low. The interview based on the matching task provided the best discriminating measure. However there was no strong confirmation of the expectation that differences between good and poor readers are larger on tasks with higher cognitive demands. It seems that measuring the awareness of text structure is complicated because several factors related to the task, the text and the reader do interact. Therefore we recommend to use at least two different measures of awareness: a recall task, and depending upon the aim of the research a composition or a matching task accompanied by an interview or a questionnaire.

References

Meyer, B.J.F. (1975). Identification of the structure of prose and its implications for the study of reading and memory. Journal of Reading Behavior, 7, 7-47.

Meyer, B.J.F. (1985). Prose analysis: purposes, procedures, and problems. In B.K. Britton, & J.B. Black (Eds.), Understanding expository text. Hillsdale, N.J.: Erlbaum.

Richgels, D.J., McGee, L.M., Lomax, R.G., & Sheard, C. (1987). Awareness of four text structures: Effects on recall of expository text. Reading Research Quarterly, 22, 177-196.

TASK CHARACTERISTICS CONTRIBUTING TO DIFFICULTY OF FUNCTIONAL READING TESTS

W. de Geus, P. Reitsma, Paedologisch Instituut Amsterdam, Duivendrecht, & K. de Glopper, Stichting Centrum voor Onderwijsonderzoek, Amsterdam, The Netherlands

Introduction

In recent years a growing interest has been noticeable in the applicability in everyday life of reading skills learned at school. In the evaluation of education this interest has resulted in the emergence of so-called functional reading tasks. The question can be asked about which components or subskills are important for a successful performance in functional reading tasks. Are different skills involved in carrying out functional reading tasks as compared with more traditional reading tasks? As a first step toward answering this question a secondary analysis is made of tasks administered in three recent Dutch reading assessments.

Kirsch & Mosenthal (1990) reanalysed the data from the NAEP literacy assessment in the U.S.A. (Kirsch & Jungeblut, 1986) in order to identify critical variables that accounted for the performance of young adults across a set of document literacy tasks. These tasks required the respondents to 'locate and use information contained in documents, such as job applications, pay-roll forms, bus schedules, maps, tables and indexes'.

First Kirsch & Mosenthal analyzed the structure of the documents and of the accompanying questions and directives by parsing them using a relational grammar, specially devised for this purpose. In this grammar three hierarchical entities were distinguished: semantic features, specific information and organizing categories.

Then they differentiated two types of variables based on this grammar, viz. *document* variables (based on the structure and complexity of the document) and *task* variables (based on the structural relation between the document and the accompanying question or directive). Apart from these, they identified three *process* variables (based on strategies used to relate information in the question or directive to information in the document). A total of 12 variables was discriminated. In a multiple regression analysis 5 of these variables accounted for 89% of the variance in the difficulty of the 61 items, as indicated by the percentage correct score (p-value). Two of the process variables alone accounted for 75% of the variance. The three other variables they found to be related to difficulty were number of specifics, number of organizing categories required by the task and number of specifics required by the task.

The present paper describes an attempt to replicate and cross-validate the findings of Kirsch & Mosenthal using the data and tests of recent large-scale Dutch assessments of literacy. The materials and p-values per item of 3 studies in which reading tasks were administered to primary and secondary school pupils were used as raw data (Wesdorp et al.,1986; Kuhlemeier & Van den Bergh,1989; Zwarts et al., 1990).
The question is whether the same variables that have been identified by Kirsch & Mosenthal as critical parameters, are necessary or sufficient to account for the variance in item difficulty in Dutch materials.

Method
An initial attempt to replicate the use of the grammar failed, because the process of parsing appeared not to be well-defined and because the grammar was not applicable to some types of tasks or formats. Moreover, it would be quite time-consuming to parse all the tasks in our dataset. Considering also the fact that grammar variables contributed relatively little to the solution of Kirsch & Mosenthal, it was decided not to use Kirsch & Mosenthal's relational grammar any further and to identify important descriptive categories or task characteristics based on relevant literature instead (e.g. Duffy, 1985; Kirsch & Jungeblut, 1986; Van den Bergh et al. , 1986; Guthrie, 1988; Kirsch & Mosenthal, 1990). The two process variables which accounted for 75% of the variance in difficulty in the Kirsch & Mosenthal study were incorporated in our variables (viz. degree of correspondence and type of information).
The aim of the study then is to determine which descriptive variables are important in accounting for differences in item difficulty.
The selected descriptive variables are at the level of the material as well as on the level of the task and the processes. Most variables were judged on five-point scales.

The 6 material-variables are:	The 8 item-variables are:
- material format - communicative function - average sentence length - relative number of difficult words - amount of text-information - organisation of the text	- item format - reference to text - amount of text-information needed - content of the question - mental activity needed - number of features - degree of correspondence - type of information

The descriptive variables were used to analyse 32 tasks (with a total of 227 items) selected from 3 national assessments of the reading performance of pupils. Two of the assessments were administered at the end of primary education (11-12 years) and one at secondary education (third grade, ±15 years). Two persons judged all items and tasks according to the 14 variables. As indicated by the data some scalepoints were quite difficult to discriminate and were therefore taken together for the rest of the analyses. A perfect agreement between raters was obtained in 75% of the assignments.

The p-values of the items were regressed on the variables in several multiple regression analyses. We entered into the regression equation those variables that had a zero-order correlation with the criterion significantly differing from zero. Analyses were done for different subsamples of items. Primary education (2 samples: PE1 and PE2) and secondary education populations (SE) were analysed separately. Table 1 shows the results of the regression analyses in the PE1, PE2 and SE samples and also includes a separate analysis of a subset of 3 tasks which were administered in both PE1 and SE.

Variable	PE1 (87 items)	SE (111 items)	PE2 (74 items)	PE1 & SE (44 items) PE1	SE
average sentence length	_	_	_	_	_
number of difficult words	_	_	_	_	_
amount of text-information	_	-.458**	.318*	_	_
organisation of the text	_	.033	.173	_	_
item format	_	-.060	_	_	_
reference to text	_	_	_	_	_
text-information needed	-.155	_	_	-.284	-.185
content of the question	_	-.073	-.095	_	_
mental activity needed	-.203	-.327**	-.063	-.140	_
number of features	_	-.057	.014	-.184	-.188
degree of correspondence	_	_	_	-.072	-.139
type of information	-.254*	_	-.120	-.073	-.209
Multiple R	.49	.55	.48	.59	.55
R^2	.24	.31	.23	.34	.31
Adj.R^2	.22	.27	.17	.26	.24

A dash indicates that the variable had an nonsignificant zero-order correlation with performance and thus was eliminated from the regression analysis for that group.
* p≤.05, ** p ≤.01

Table 1
Standard Beta coefficients of predictive variables in 5 samples

Results

A different set of predictors for item difficulty was found for each sample (Table 1). This variable set of predictors could either be the result of differences in sample population or of differences in tasks included. We checked this by separately analyzing a subset of 3 tasks (44 items) which were administered in both PE1 and SE. The results show that over the two subpopulations the same variables accounted for roughly the same amount of explained variance in p-value in these tasks (Table 1). This finding suggests that differences in tasks included are far more important than differences in sample population in influencing the prediction of item-difficulty in our data.

Kirsch & Mosenthal's analyses were done on document literacy tasks: i.e. no prose tasks were involved. To be able to compare our results to theirs, we did separate regression analyses on document-like and prose-like tasks per sample. We also

did some further analyses with other subsets of our data to explore the effects of multiple choice vs. open ended questions and more difficult vs. easier items, without any consistent findings, however.

Conclusions

The Kirsch & Mosenthal results were not replicated. Not only is the total amount of variance accounted for in our study less, but also the amount of variance accounted for by the process variables we took from their study. Moreover, dependent on the sample of tasks we found different variables to be related to performance. A constant finding in most analyses was the amount of variance accounted for; in most samples it was limited to 30%. Considering the number of predictors this is a meagre result, especially when compared to the Kirsch Mosenthal results. Yet it should be mentioned that only 56% of variance in the p-values in the least educated group of respondents (0-8 years of education) which is likely most comparable to our samples was accounted for in the Kirsch & Mosenthal study and only by one variable, namely type of information.
It is also possible that the Kirsch & Mosenthal findings are attributable to a fortuitous coincidence of items. According to our results the composite of tasks that is analysed is highly determining for the predictability of the difficulty of the items. On top of that taskmembership as a dummy-variable happened to be an important predictor for item-difficulty in two of our samples. Exactly what is common to items belonging to one task is not explained by our task variables.

References

Bergh, H. van den, Baltzer, J. E., Marschall, F. T., & Wesdorp, H. (1986). Instrumenten ter peiling van de taalvaardigheid van leerlingen uit de derde klas V.O. (SCO-rapport nr. 71). Amsterdam: SCO.

Duffy, T. M. (1985). Readability formulas: what's the use? In: T. M. Duffy, & R. Waller (Eds.), Designing usable texts. London: Academic Press.

Guthrie, J. T. (1988). Locating information in documents: examination of a cognitive model. Reading Research Quarterly, 23(2), 178-199.

Kirsch, I., & Jungeblut, A. (1986). Literacy: Profiles of America's Young Adults. (Report no. 16-PL-02). Princeton, N.J.: Eductional Testing Service.

Kirsch, I. S., & Mosenthal, P. B. (1990). Exploring document literacy: variables underlying the performance of young adults. Reading Research Quarterly, 25(1), 5-30.

Kuhlemeier, J. B., & Van den Bergh, H. (1989). De Proefpeiling Nederlands: een onderzoek naar de haalbaarheid van peilingsonderzoek in het voortgezet onderwijs. (Specialistisch bulletin nr.74). Arnhem: Cito.

Wesdorp, H., Van den Bergh, H., Bos, D. J., Hoeksma, J. B., Oostdam, R. J., Scheerens, J., & Triesscheijn, B. (1986). De haalbaarheid van periodiek peilingsonderzoek. Een voorstudie op het gebied van het taalonderwijs in de lagere school. (SVO-reeks 85). Lisse: Swets & Zeitlinger.

Zwarts, M., Janssens, F., Rijlaarsdam, G., Veldhuijzen, N., Wesdorp, H., & Wolfhagen, I. (1990). Balans van het taalonderwijs aan het einde van de basisschool. Uitkomsten van de eerste taalpeiling einde basisonderwijs. (PPON-reeks nr. 2). Arnhem: Cito.

QUALITY OF HANDWRITING: SELF-ASSESSMENT AS A TOOL FOR IMPROVEMENT

M. Boekaerts, C. Nieuwenhuizen, & G. Seegers, Centre for the Study of Education and Instruction, Leiden University, Leiden, The Netherlands

Meulenbroek and Van Galen (1988) refer to mature handwriting as medium-speed, ballistic movements. They argue that the development of mature handwriting strategies is mostly completed at the age of fifteen. In the initial stages of learning the skill pupils learn to behave according to strict rules of a specific handwriting curriculum, but gradually they develop their own handwriting skill as a means to express their thoughts and feelings to others. In traditional handwriting lessons teachers instruct pupils how to write letters, words and sentences. Meulenbroek (1989) identified three stages: viz. a preparatory stage, a primary stage and an advanced stage. In the *preparatory* stage global training in perceptual-motor skills is given and the child is made familiar with the written symbols of the language, and given a lot of drawing tasks. In the *primary* stage, teachers introduce students to so-called 'methodological handwriting'. The writing method they use prescribes form, slant, and size of the letters, as well as how the connections between letters should be made. These writing rules are demonstrated in the teachers' handwriting, and they are also displayed in the master script table on the classroom wall, and in the pupils' workbooks. Students have to adopt these rules and practise them, even though their parents and siblings may display different forms of handwriting. *Advanced* handwriting starts in the 3rd grade (around the age of eight). At that point in the curriculum, writing methods introduce exercises in which increasingly larger pieces of text have to be written. Teachers are requested to devote more attention to the speed of writing, and to allow their pupils to develop their personal writing style. At the same time, they are urged to inspect the regularity of the pupils' handwriting, and its overall legibility.

It is important to note, that in the advanced stage most pupils are able to copy and spontaneously form correct letters and words, but that their writing performance may nevertheless show irregularities. When irregular writing patterns (e.g. inconsistency in letter formation, letter size, slant, and letter/word spacing) are tolerated, pupils may develop immature handwriting strategies. Indeed, as several authors displayed, the main difference between good and poor writers is that the latter display form variability and discontinuity of strokes (cf. Mojet, 1989; Wann & Kardirkamanathan, 1991).

Feedback and self-assessment

The central question we wanted to answer was: 'What can be done to promote consistent and regular handwriting? There are in our opinion two crucial mechanisms for the development and maintenance of any psycho-motor skill: viz.

- observation of adequate, dynamic model examples, and
- adequate feedback..

When students are trained in the primary stage, they watch their teacher while writing on the blackboard. Such observations provide them with a dynamic model. However, when they start copying the examples written on the blackboard, or in their copybooks, the written letters, words, and sentences have become static products of the handwriting act. This means that pupils have to remember how to form and connect the different letters. Encoding information into motor memory and activating it from motor memory, are complex processes described in great detail by Thomassen and Van Galen and their co-workers. For example, Meulenbroek (1989) pointed out that pupils have to build up stable internal representations (motor programs) through a lot of practice, and that explicit training is necessary to ensure swift retrieval of motor programs.

Apart from practice, feedback is also crucial for the development of a smooth on-line programming strategy. Feedback can be immediate (on-line) or postponed (retrospective) in nature. When pupils are copying examples from the blackboard or from model sentences in their workbooks, the teacher can walk around in class and provide fast and accurate feedback while watching the pupil write (on-line feedback). Alternatively, the teacher can comment on finished products (retrospective feedback). The latter type of feedback will concentrate mainly on spatial aspects such as size, slant, width, lineation. The former type of feedback may also include comments on dynamic characteristics of the handwriting performance, such as e.g. velocity, acceleration, fluency of movement, pen pressure, writing position etc. Results from experimental investigations on writing performance, indicate that various forms of feedback are differentially effective in the different stages of the writing acquisition process. For example, a study by Portier and Van Galen (in press), in which new writing patterns were taught to adults, displayed that students who received postponed feedback produced more fluent scripts than students who were provided with immediate feedback. But, in the immediate feedback condition there was stronger evidence for a shift toward an on-line programming strategy. They concluded that immediate feedback results in better learning performance, whereas postponed feedback makes the writing skill more fluent.

These results seem to imply, that in the primary stage of the handwriting acquisition process, it is better to provide on-line feedback about a pupil's performance, whereas in the advanced stage postponed feedback is sufficient to increase fluency. These results are in line with our theorizing, that in the primary stage, or in remedial work, immediate feedback from the teacher is a crucial aspect of the writing acquisition process. It should be given about both spatial and kinematic aspects. However, once pupils have acquired the basic motor programmes and can activate them swiftly, they can be trained to provide their own feedback by learning to judge the quality of their own handwriting. In such self-assessment training, pupils should be provided with criteria, and their attention

should explicitly be drawn to irregularities in the handwriting process. It was hypothesized that such self-asseeement training would enable third graders to develop an internal standard of 'good handwriting', and that it would improve the quality of their handwriting, as well as their attitude toward writing. In addition, this skill would make pupils relatively independent of the teacher.

Constructing a self-assessment scale

Bruinsma and Nieuwenhuis (1991) pointed out that teachers use subjective ratings to assess their pupils' performance, and that they find it difficult to verbalize the criteria they use. These authors constructed a scale in which five basic criteria were used to judge the quality of handwriting: viz.

- *general appearance*, or the rater's overall impression of the quality of the student's writing;
- *slant*, or the angle formed by the direction of the down strokes with respect to the bottom line;
- *form*, or matching individual letters with the standard used in the handwriting lesson;
- *lineation/space*, or the distance used between letters and words, and
- *size*, or the height of letters and the ratio of height and width.

They reported that student teachers can learn to apply these criteria reliably. The aim of the present study was to investigate whether pupils, who are either in, or on the brink of the advanced stage of writing development, could be trained to use these five criteria to assess their own, and their peers' handwriting. We also wanted to know whether self-assessment skill would improve their writing performance, and would contribute to a more positive attitude towards writing in class. A training programme was developed in which an experimental and a control group of third graders (average age around eight) participated during five consecutive weeks (half an hour practice a day). At the beginning and end of the training programme the quality of the pupils' handwriting was measured, as well as their attitude towards writing. The results of the self-assessment training will be discussed.

References

Bruinsma, C., & Nieuwenhuis, C.W. (1991). Criteria for the evaluation of handwriting. An empirical study on the effects of training student teachers in the use of evaluation criteria. Paper presented at the 5th Handwriting Conference of the International Graphonomic Society, Arizano, October 27-30.

Meulenbroek, R. (1989). A study of handwriting production: Educational and developmental aspects (Doctoral dissertation). Nijmegen: NICI.

Meulenbroek, R.G.J., & Galen, G.P. van. (1988). The acquisition of skilled handwriting: Discontinuous trends in kinematic variables. In A.M. Colley & J.R. Beech (Eds.), Cognition and action in skilled behaviour (pp. 273-281). Amsterdam: North Holland.

Mojet, J.W. (1989). Characterisitcs of the developing handwriting skill in elementary education. In J. Wann, A.M. Wing & N. Sovik (Eds.), Development of graphic skills. Research perspectives and educational implication. London: Academic Press.

Portier, S.J., & Galen, G.P. van. (in press). Immediate versus postponed visual feedback in practising a handwriting task. Human Movement Science.

Wann J.P., & Kardirkamanathan, M. (1991). Variability in children's handwriting: Comparative diagnosis of writing differences. In J. Wann, A.M. Wing & N. Sovik (Eds.), Development of graphic skills. Research perspectives and educational implication (pp. 223-236). London: Academic Press.

FUNCTIONAL WRITING ASSIGNMENTS AND KNOWLEDGE ACQUISITION

G.M.T. Poorthuis, & W.A.M. Kok, ISOR, Department of Educational Research, University of Utrecht, Utrecht, The Netherlands

Introduction

Functional writing assignments in academic subjects can induce pupils to process textbook materials actively and independently (Newell, 1984; Durst & Newell, 1989). Higher cognitive activities can be initiated by means of such assignments, which in turn may lead to concept formation and the structuring of knowledge. Present information processing theories stress that the intensity and quality of the cognitive activities of the student are co–determinants of the quality of acquired knowledge (Anderson, 1983; Anderson & Reder, 1979; Mayer, 1980). Cognitive activities which are a result of functional writing assignments in learning a subject can thus lead to improved knowledge of that subject.

There is a distinction between open and guided assignments in functional writing assignments. In open writing assignments, the pupil is free in terms of the design and execution of the writing process. In guided writing assignments the design and execution of the writing product are prestructured by the assignment itself.

In open writing assignments low composing ability will cause the pupil to use a "knowledge telling strategy" (cf. Andriessen & Boonman, 1988; Bereiter & Scardamalia, 1987; Flower, 1981; Flower & Hayes, 1980). This type of writing strategy will contribute relatively little to the acquisition of knowledge by the pupil. Pupils with low composing ability will benefit more from guided writing assignments, where the writing task is prestructured and the assignment initiates cognitive activities which contribute to concept formation. Pupils of high composing ability will generate such higher cognitive activities also in open writing assignments, since they make use of a 'knowledge transforming strategy'.

The present study examines the relationship between functional writing assignments and knowledge acquisition in geography and physics in secondary education. In order to demonstrate the effects of functional writing on learning, the following research questions were addressed:

- What is the direct and the long–term effect of the various types of writing assignments within a school subject on pupils' acquired knowledge?
- Do the various types of functional writing assignments have a differential effect on pupils' acquired knowledge as a result of differences in composing ability?

ERIC

- What is the effect of the quality of the product of the functional writing assignment on acquired knowledge?

Research method

Design

Conditions:

exp.group 1	n ± 30	o1	o2	x1	o3	o4	o5
exp.group 2	n ± 30	o1	o2	x2	o3	o4	o5
control group	n ± 30	o1		y		o4	o5

o1 = assesment of pre-existing knowledge
o2 = measurement of composing ability
o3 = evaluation quality of the product of writing assignment
o4 = direct measurement of acquired knowledge
o5 = delayed measurement of acquired knowledge (durability)

x1 = open writing assignments
x2 = guided writing assignments
y = study assignment without writing task

In accordance with this design, three measurements were carried out in October, January and April 1989–1990.

Students

The study involved the subjects of geography and physics and was carried out at a MAVO school (lowest form of secondary education) and a HAVO/VWO school (higher forms of secondary education). For each phase of the study, three parallel classes participated in which instruction in geography and physics, as well in Dutch language, was given by the same subject teacher. The content and method of teaching in the three conditions was therefore controlled as much as possible. Approximately 90 students, divided over the three conditions, participated in each of the subjects.

Treatment

During the school year of 1989/1990 the students in the experimental conditions were regularly given writing assignments while in the process of studying the subjects mentioned. Feedback sessions were used to effect improvements in the writing process.

Instruments

The composing ability of the students was measured by judging the quality of written essays with the use of an analytical evaluation procedure. In order to evaluate the acquired knowledge tests were developed for each school type, subject and time of measurement. The quality of the texts produced as a result of the writing assignments was measured in terms of the number of concepts and relations from the original instruction text that students produced in their texts.

Results

First research question

By means of a covariance analysis we established whether the mean knowledge–scores of the three experimental groups differed significantly from each other. This was done separately for each school type, subject and time of measurement. Preknowledge was included as a co–variable.

Only HAVO/VWO students showed any significant differences, in the subject of geography. The students who had worked with open writing assignments turned out, on average, to have significantly higher scores in the knowledge tests than students without writing assignments (i.e. the control group) and sometimes also higher than students who had studied the subject with the help of guided writing assignments. In a few cases the knowledge–test scores for the latter group were higher than those for the control group.

It may be concluded that in geography the use of open writing assignments by HAVO/VWO students can result in improved knowledge. In physics, however, not a single trace of evidence was found that writing assignments have any effect, either positive or negative, on acquired knowledge.

Second research question

With the use of analyses of covariance, we tested the interaction between types of writing assignment (open vs. guided) and composing ability (low, medium, high). A significant interaction effect in geography was found for HAVO/VWO, in the context of some delayed measurements. Pupils with low composing ability who had worked with open writing assignments showed higher scores on the knowledge test than pupils with an average or high composing ability. By contrast, in guided writing assignments pupils with low composing ability scored higher than pupils with high composing ability.

On further analysis it turned out that, with regard to composing ability, this interaction effect was due to the aspects of 'structure' and 'content'. Both these aspects of composing ability were found to interact with type of writing assignment, with the effects of the interacting variables operating in the same way as for composing ability in general.

The conclusion is that open writing assignments in geography can have a positive effect on acquired knowledge, in particular for HAVO/VWO pupils with low composing ability. Again, no detectable effect at all was found for physics.

Third research question

The question about the effect on acquired knowledge of the quality of the product of the functional writing assignment was answered by means of regression analyses, using quality of the product, together with preknowledge and experimental condition, as independent variables and scores on the knowledge test as dependent variable. The contribution of the quality of the produced text to the explained variance in scores on the knowledge test practically always turned out to be non–significant and invariably very small.

The quality of the writing product is also immaterial in those cases where the completion of open writing assignments leads to improved knowledge of the subject. This might indicate that it is not the writing product that influences the quality of acquired knowledge so much as the writing process itself.

320

References

Anderson, J.R. (1983). The architecture of cognition. Cambridge, M.A.: Harvard University Press.

Anderson, J.R., & L.M. Reder (1979). An Elaborative Processing Explanation of Depth of Processing. In: L.S. Cermak & F.I.M. Craik (Eds.), Levels of Processing in Human Memory. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Andriessen, J.E.B ,& J.H. Boonman (1988). Probleemoplossen en het begrijpen en produceren van teksten, Pedagogische Studiën, 65, 1, p. 32–44.

Bereiter, C., & M. Scardamalia (1987). The Psychology of Written Composition. Hillsdale, N.J.: Erlbaum.

Durst, R.K., & Newell, G.E. (1989). The Uses of Function: James Britton's Category System and Research on Writing. Review of Educational Research, 59, No. 4, 375–394.

Flower, L.F. (1981). Problem solving strategies for writing. New York: Hardcourt, Brace & Jovanovich.

Flower, L.F., & J.R. Hayes (1980). The dynamics of composing: making plans and juggling constraints. In: L. Gregg & E. Steinberg (Eds.) Cognitive Processes in Writing: An interdisciplinary approach. Hillsdale, New Jersey: LEA.

Mayer, R.E. (1980). Elaboration Techniques that Increase the Meaningfulness of Technical Text: An Test of the Learning Strategy Hypothesis. Journal of Educational Psychology, 72, 770-784.

Newell, G.E. (1984). Learning from Writing in Two Content Areas: A Case Study/Protocol Analysis. Research in the Teaching of English, 18, No. 3, 265–287.

ACTIVE KNOWLEDGE OF NUMBERS IS LIMITED AND CONTEXT-BOUND

M. Milikowski, & J.J. Elshout, Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

Abstract

In this paper we examine the results of four experiments, arguing that the active numerical vocabulary of adults is limited to the numbers 1-13 with the addition of some multiples of ten and eleven. It is further argued that within this small set of actively known numbers two factors operate to determine a number's chance of being actively used: a domain specific and and domain general one. Numbers like 7 and 13, which attain no prominence within a numerical association context, are decidedly favoured in a more general context. The opposite is true of numbers like 2 and 4, which gain prominence of access in a context of numerical association, but are much less popular in general.

Summary

The objective of our first experiment was to determine 'meaningfulness' scores for all natural numbers up to 100. The 'meaningfulness' or 'm'-score of a concept is defined as the number of different associations it will generate during (half) a minute, averaged over subects (Noble 1963). An interesting discovery made while studying the results concernes the narrowness of the range of numerical responses; while all numbers from 1 through 100 were presented as stimuli equally often, their probability of being named as a response was distributed most unevenly. To illustrate: while the number 2 , a general favourite in this experiment, is used as a response with a total frequency of 533 times (which amounts to nearly thirty times per subject) the number 79 is used as a response only 20 times (which is little more than once per subject). A number's frequency of being used as a response we called its 'p'-score (Elshout and Milikowski, 1990). We wish to argue here that the p-score of a concept is a measure of its availability as a piece of active knowledge.

In our second and third experiments, which aimed to explore semantic links between numbers, 104 subjects were asked to name just one number in response to a given one. Again all numbers from 1 through 100 were presented as stimuli equally often; again the response pattern showed a distribution as markedly uneven as the pattern obtained in experiment 1, p-scores in the different experiments correlating .90 and more. In all three experiments, the numbers 1 to 10 account for more than 50 percent of the responses. Some numbers between

10 and 25 are quite often mentioned, as are the smaller multiples of ten and eleven. Except for those, numbers above 25 are seldomly named.
Our claim that the phenomenon described above is not an artefact of the tasks in question is supported by the results of the fourth experiment,. in which more than 400 subjects produced twenty numbers each, in response to a series of verbal cues..Subjects were invited to name a) their favourite numer; b) a number they disliked more than any other; c) three pleasant and d) three unpleasant numbers; e) three interesting and f) three uninteresting numbers; g) three calm and h) three excitable numbers. Subjects readily performed this task, but again predominantly used the numbers 1-13 as responses. It appears, in sum, that the active numerical vocabulary of university freshmen is limited to a very small set of numbers.
Within this set of well known numbers, two factors seem to operate in determining relative precedence. The first factor seems to be concerned with the "helpfulness" of numbers in the context of everyday arithmetics, high scoring numbers being 2, 4 and 10. The other factor is associated with prominence in non-mathematical contexts, 3, 7 and 13 gaining leading positions.

References

Elshout, J.J., & Milikowski, M. (1990). De produktie van geautomatiseerde rekenkennis. In M. Boekaerts en E. de Corte (eds). Onderwijsleerprocessen, ORD 1990. Nijmegen: ITS

Noble, C.E. (1963). Meaningfulness and familiarity. In C.N Cofer and B.S. Musgrave eds. Verbal behavior and learning, problems and processes. New York: McGraw-Hill.

A PILOT STUDY ON THE POSSIBILITIES OF REALISTIC MATHEMATICS INSTRUCTION IN SPECIAL EDUCATION

Marja van den Heuvel-Panhuizen, Freudenthal Institute, University of Utrecht, Utrecht, The Netherlands

Introduction to the problem

There is a wide gap in The Netherlands between mathematics instruction in regular primary schools and schools for special education. The innovation of mathematics instruction as it occurred in the field of regular education (Treffers, 1987; De Jong, 1986) did not find acceptance in special education. In contrast with the realistic approach of mathematics in regular primary schools, the approach in schools for special education is still rather mechanistic. Although arguments which plead for a more realistic approach of mathematics in special education (Van den Heuvel-Panhuizen, 1986; Ter heege, 1988) have not been completely without effect, in general both teachers and psychologists in special education remain most reluctant to shift to a more realistic approach. Besides uncertainty which results from a lack of research data, objections arise in respect to group instruction, the building on informal knowledge of the children and its connected variety of solution strategies, and because of the interaction in the classroom, and the complex factor of starting from contexts (Van Luit, 1987 and 1988; Damen, 1990). The aim of the present study is to gather some data on the possibilities of realistic mathematics instruction in special education.

Research method

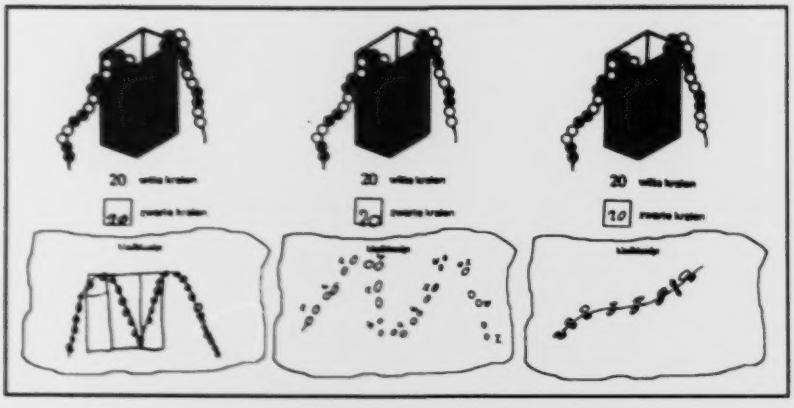
The best way to do this would be to carry out a teaching experiment with an experimental group which is taught in the realistic manner. However, by lack of means, here a choice has been made for a different approach. Data on the possibilities of realistic instruction have been gathered without involving any kind of realistic instruction, even without any instruction at all. The data have been gathered with the aid of a test only. In order to test the attainability of realistic mathematics in a difficult situation, the study was directed at mentally retarded children. Moreover, a topic has been chosen which is not regularly included in the mathematics program of schools for mentally retarded children (Thornton et al., 1983; Damen, 1990): ratio.

The test on ratio which has been developed was devised in accordance to the manner this topic is dealt with in realistic mathematics instruction (which contrasts markedly with the mechanistic approach of it) and to the principles of the MORE-tests (Van den Heuvel-Panhuizen, 1990; Van den Heuvel-Panhuizen &

Gravemeijer, 1991). The first implies that the test contains a variety of different situations in which children encounter ratio and which they are somehow familiar with. The latter means that the children are offered opportunities to show optimally what they are capable of. The test consists of sixteen items and contains four different kind of ratio problems: finding the ratio, comparing ratios, producing equivalent ratios, and finding the fourth proportional.
The test was administered to 61 children from the two upper grades (grade 5 and 6) of two schools for mentally retarded children. Besides the pupils' capability of solving ratio problems other pupil data have been collected like age, sex and math level. Also an inventory was made of the math topics that had been dealt with. Further more, the four teachers of the pupils and two inspectors and two educational psychologists for special education have been asked to make an estimate in advance per test item of the number of pupils that would do the pertaining test item correctly.

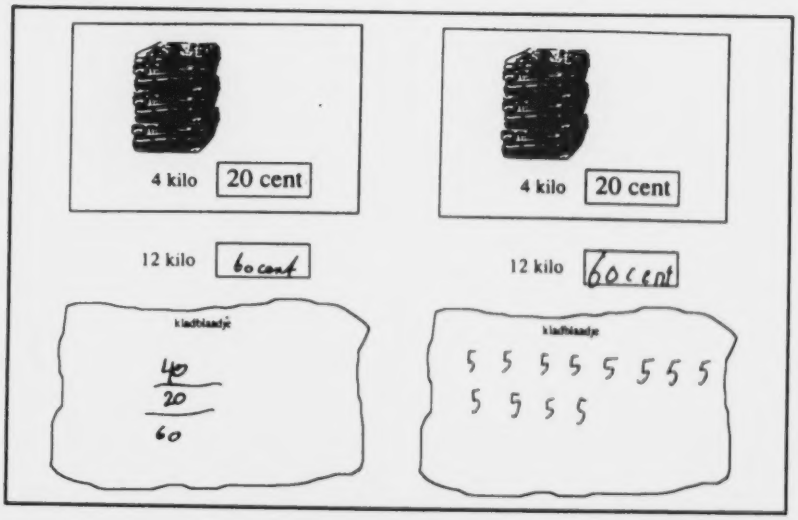
Results
The test has a reasonable internal homogeneity (alpha = 0.61) and the frequency distribution of the pupils' total scores is rather normal without top and bottom effect. The percentages correct of the items lie between 13% and 64%. Half the items have a percentage correct of more than 40%. The scrap papers placed on the test sheets reveal how the children arrived at their answers. See for instance figure 1. Given in this test item (no. 12) is the number of white beads and the children are asked how many black beads the chain has. The problem is that the children cannot see all of the beads. Half of the children, however, came to the correct answer here The pieces of scrap paper in figure 1 show the use of a model at three different levels to arrive at the solution. The most concrete one is the one on the left, the most abstract the one on the right. In the latter case both the number of beads and the specific pattern of two black beads followed by two white beads are no longer important. The only thing that counts is that there is an equivalent relationship between the black and the white beads.

Fig.1



A more difficult item (no. 15) is showed in figure 2. Here the children have to calculate the price of twelve kilos of newspaper. A quarter of the children were capable of solving this problem. Again the scrap papers show how. The child of the left one arrived at the answer by means of three times four kilo. The child of the right one first calculated the price per kilo.

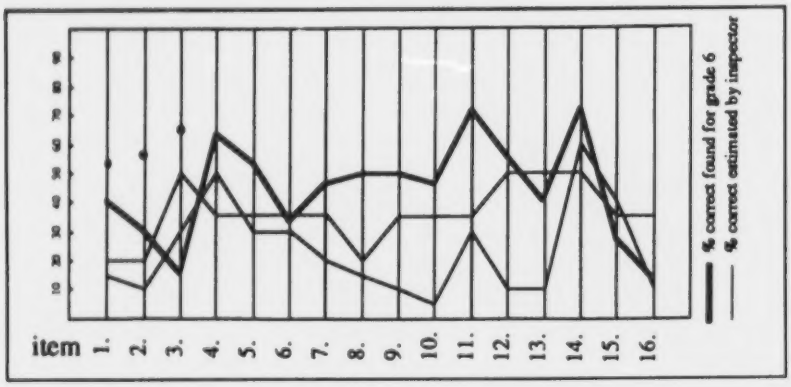
Fig.2



These test results contrast sharply with the results of the inventory of the contents of the mathematics program: in both schools no instruction in the area of ratio took place.

The estimates on the percentages correct showed that the skills of the pupils at the end of primary school for mentally retarded children were underestimated on a fair number of points. For the inspectors (see figure 3) this applied more than for the educational psychologists, but for the two grade 6 teachers this is only the case for one of them.

Fig.3



As regards the investigation of the relationship between test score and certain pupil characteristics only the connection with math level in the class appeared to be significant.

Conclusions

The test results indicate that a topic like ratio, has unjustly been left out of the mathematics program in special education. Further, the evidence on the scrap papers illustrates that inventing and using strategies and reporting on them, is anything but impossible for pupils at a school for mentally retarded children. Moreover the test results reveal that contexts do not have to be the limiting factor they are so often thought to be in special education.

Discussion

Although no general conclusions can really be made on the basis of this limited study these findings point out that there is at least reason to subject the mathematics program of special education and the didactics involved. As such this study affords some cues for innovation. The special feature of this study is that these cues have been provided by means of a class administered test. As a matter of fact, this involves a new role of tests in innovation. Tests can contribute towards improving education. A prerequisite for this should be, however, that tests, among others, offers children opportunities to show what they are capable of.

References

Damen, S. (1990). Rekenen op lom- en mlk-scholen. Ervaringen tijdens een stage bij het Instituut voor Toetsontwikkeling (CITO). Arhnem: Cito.

Treffers, A. (1987). Three dimensions. A model of goal and theory description in Mathematics Instruction - the Wiskobas Project. Dordrecht: Reidel.
De Jong, R.A. (1986). Wiskobas in Methoden. (Dissertatie). Utrecht: OW&OC.
Heuvel-Panhuizen, M. van den (1986). Het rekenonderwijs op de lom-school opnieuw ter discussie. Tijdschrift voor Orthopedagogiek, 25 (3), 137-145.
Heuvel-Panhuizen, M. van den (1990). Realistic Arithmetic/Mathematics Instruction and Tests. In K. Gravemeijer, M. van den Heuvel & L. Streefland. Contexts. Free Productions. Tests and Geometry in Realistic Mathematics Education. Utrecht: OW&OC.
Heuvel-Panhuizen, M. van den, & Gravemeijer, K.P.E. (1991). Tests are not all bad. In L. Streefland (ed.) Realistic Mathematics Education in Primary School. Utrecht: Freudenthal Institute.
Heege, H. ter (1988). Realistisch reken-wiskundeonderwijs. Op weg naar vernieuwing in het speciaal onderwijs. School & Begeleiding, 5 (17), 12-14.
Luit, J.E.H. van (1987). Rekenproblemen in het speciaal onderwijs. (Dissertatie). Nijmegen: Katholieke Universiteit Nijmegen.
Luit, J.E.H. van (1988). Realistisch reken-wiskundeonderwijs in het speciaal onderwijs? School & Begeleiding, 5 (17), 15-18.
Thornton, A., Tucker, B.F.,Dossey, J.A., & Bazik, E.F. (1983). Teaching Mathematics to Children with Special Needs. Menlo Park, California: Addison Wesley Publishing Company.

STUDY ORIENTATIONS AND LEARNING IN MATHEMATICS BY UPPER-SECONDARY SCHOOL STUDENTS

Raija Yrjönsuuri, University of Helsinki, Helsinki, Finland

The presentation describes a socio-cognitive theory of orientation and its application to mathematics learning in upper secondary schools. Four study orientations are defined: problem solving, social dependence, ego-defensive and giving-up (abandoning) study orientation. This presentation, which is based on two inquiries (1989, 1990), deals with the experiences, expectations, activities and learning of students in the survey of mathematics.

The experience of meaningfulness in school is by its nature greatly a product of social interaction.The learning process cannot be described as a passive adoption of some structure of knowledge that is given from outside, but as an active process, where the individual is in continuous interaction with his environment. Learning is represented according to the constructivistic view. Study, as an element situated between teaching and learning, adds to the picture the traits of the students self-concept as a learner. Which situational factors does the student give greatest weight to in his orientation towards a learning task?

The activity of a problem solving oriented student is directed by inner motivation and hope for success towards the solution of the task and the acquisition of knowledge that is required for it. The learning process is characterized by an orientation towards the future.

A student that is characterized by social dependence study orientation concentrates on social environment of the moment. Performance processes get positive meaning by being important to persons that the student esteems. The student concentrates on the stage of instruction at hand. The study process has instrumental meaning for the student.

A student with ego-defensive study orientation connects the causes or results of earlier failures to external reasons, for instance the difficulty of the task, rather than to his own endeavours. If a student associates failure with his inability, it leads to a growing feeling of inferiority and to further failures. The past is predominating in the anticipation of the future. The ego-defensive study orientation is manifested in many kinds of hindrances of learning like the difficulty of mastering mathematical concepts of perceiving the problem, lacking self-confidence. (Olkinuora 1988).

A student characterized by giving-up study orientation does not experience the study of mathematics as meaningful to himself, but he can experience the study of some other subject as highly meaningful.

Study orientations and learning mathematics are regarded as phenomena which have relation to each other. The classifications of mathematical problems are an application of the so called Wilson model.

The correlations between perceived achievements and notes in mathematics were statistically significant. The investigated situational orientations were typical for the mathematics learning. Common characteristics were e.g. a weak problem solving orientation together with high operational activeness, a negative problem solving orientation together with learned helplessness, a high social dependence orientation together with high operational activeness, a high ego-defensive orientation together with difficulties in understanding mathematical language, texts, and problems, and a high abandoning orientation together with low interest in learning mathematics. The ego-defensive and the social dependence orientations illuminated and interpreted students' difficulties in learning mathematics.

The four main characteristics of learning mathematics were use of symbols and operations, knowing concepts, generalizing mathematical structures, and testing problem solving. All four are hierarchical: the first and second can be interpreted as using cognitive skills; the third and fourth can be interpreted as reflective thinking. Using cognitive skills has stronger relationship with the note in mathematics than the reflective thinking has. The best explainers for the notes in mathematics were found to be problem solving orientation, specifying mathematical language, reflective and generalizing thinking, knowing meaning of symbols, and knowing relations between symbols and operations.

About one half of those who participate in the maturity exam of mathematics in the short course and about one fourth of those who do not participate are successful in learning mathematical structures and concepts. When curricula are being made, it is assumed that students have a positive attitude towards learning mathematics and that they experience learning it as meaningful. The learning process has, however, often instrumental meaning for a student. It is not anticipating learning but the expectations of the social community. The student is directed by the social need of being accepted by the community and fulfilling its expectations. Learning difficulties arise, as the student lacks confidence in his own ability to learn. Performance processes get a positive meaning by being important to persons that the student esteems highly. As a result of failures the student begins to become estranged from his responsibility for learning and ends up in learned helplessness or abandoning study. Do the goals of instruction and their realization go in different directions?

The central task of the teacher is the maintenance of social interaction and positive atmosphere in the classroom cummunity by being interesting in every student as an individual. The student underlines 'the teacher's professional skill', which he interprets as social psychological and educational interaction. The student trusts the teacher's cognitive mastery of the subject, if he only 'was able and patient enougt to give the students time to learn'. The central reason for the arising of ego-defence is the disproportion of the time available and the contents to be learnt as well as the proportioning of the whole of upper secondary school to the extramural interests of young people, which also shall have an essential task in development of a young person.

References

Olkinuora, E. (1988). Pupil's orientations, school achievements and sense of control: A follow-up study of adaptive development. Nordisk Pedagogik , 8 (2), 59-69.

Yrjönsuuri, R. (1989). Lukiolaisten opiskeluorientaatiot ja menestyminen matematiikassa. (Upper secondary school students' study orientations and school achievements in mathematics.) Helsingin yliopiston kasvatustieteen laitos. Tutkimuksia 120.

Yrjönsuuri, R. (1990). Lukiolaisten matemaattisen ajattelun oppiminen. (Upper secondary school students' learning mathematical thinking.) Helsingin yliopiston opettajankoulutuslaitoksen tutkimuksia 88.

PROOF GENERATION IN GEOMETRY: A COMPUTER SIMULATION OF THINK-ALOUD PROTOCOLS

M. Anselem, Paris 8 University, F. Levy, L.I.P.N., & A. Nguyen-Xuan, C.N.R.S. and Paris 8 University, France

The experiment that was conducted is an experiment of learning by solving problems. During ten individual one-hour sessions, three 12-to-13-year-old third-grade secondary French school students (12-13 year old) were asked to verbally demonstrate Geometry problems. In the French syllabus, proof production is introduced only in the third grade. Consequently the subjects were beginners in conducting demonstrations. They were allowed (and prompted) to use a list of "theorems" which comprised the theorems and definitions which had been previously studied in the classroom. The experimenter intervened only to refute incorrect knowledge (by offering counter-examples) when the subject used such knowledge to draw a conclusion. For instance, one subject said "segment AB is equal to segment DE, because they are the third sides of two triangles (triangles ABC and DEF) the two first sides of which are equal" (BC=EF, and CA=FD). The experimenter showed the student a figure where two triangles had two sides respectively equal and their third side unequal.

A computer model was built that simulates the recorded verbal protocols. Our immediate objective aimed at evaluating the students' progress during the ten sessions. A further objective was to elaborate a general framework to analyze verbal protocols.

We postulate that the student's knowledge can be analyzed into three components: strategic knowledge for using deduction rules in the process of proof generation; linguistic knowledge for interpreting the problem statements, the questions and the theorems stated in natural language; and intuitive knowledge in Geometry that is used in the process of linguistic interpretation and in the process of proof generation. Therefore the system that simulates the observed (verbal) behavior possesses three modules of knowledge corresponding to these three kinds of knowledge, together with a semantic network for representing the concepts and their organization. The semantic network comprises two categories of concepts: the linguistic concepts for comprehending the natural language used in the domain and the geometry concepts. The properties of the geometry concepts are correct but they are not necessarily complete. They represent third-grade students' knowledge of Geometry. For instance, a square is not explicitly described as a sub-class of a rectangle, or a sub-class of a rhomb.

The linguistic module uses the semantic network to build facts (geometric objects and final goals) from the problem statement and the questions that are given in

natural language. It creates a representation of the theorems chosen by the strategic module. The way the theorems are interpreted depends on the knowledge about geometry concepts of the semantic network. The intuitive knowledge module contains the student's basic knowledge in Geometry that are not given in the theorems. There are rules for creating the vertices and the sides of a polygon, for giving a name to a geometric object, for conducting simple deductions such as "If two segments share two points, then they are on the same line". There are also incorrect rules that have been used by the subject for drawing conclusions. These incorrect rules will be removed when the experimenter presents the counter-examples.

The strategic module is an expert-system that can perform backward and forward-chaining. This module monitors the process of proof production: it tries first to generate a proof using backward chaining search. All the theorems and definitions are represented as production rules. The search is based on the instantiation of the right-hand side of a theorem by the conclusion to be proved. In case more than one theorem has its right-hand side instantiable by the conclusion, the strategic module uses meta-rules for choosing among them. The theorem chosen by the strategic module is passed to the linguistic module for creating an internal representation which may not be a mathematically correct representation of the theorem that has been read. If a theorem exists the right-hand side of which is instantiated by the conclusion to be proved, the proof is achieved whenever the entire theorem is instantiated. Therefore, the subgoals of the search consists in trying to instantiate the left hand-side of the theorem.

If no theorem exists the right-hand side of which can be instantiated by the conclusion to be proved, or if this search method fails, the system switches to the forward chaining approach. It tries to use the rules of the intuitive knowledge module to deduce a maximun number of new facts. The selection of the deduction rules of the intuitive knowledge module is based on the "context". We define formally two different kinds of context: "geometric figure context", and "relation context". The contexts are calculated on the basis of the relation (or on the basis of the geometric object) which the conclusion to be proved refers to. The newly deduced facts may instantiate the conclusion to be proved: the demonstration is achieved. If this is not the case, the strategic module will use these facts, together with the old facts, to instantiate (via the linguistic module) the left-hand side of a theorem in the list of theorems. The proof is achieved when there is a theorem that can be entirely instantiated, with its right-hand side instantiated by the conclusion to be proved.

In conclusion, first, the model we present takes into account the students' linguistic knowledge for comprehending natural language and the mixed search strategy usually observed in beginners. Second, we separate the student's basic (mathematic) knowledge in Geometry from the "verbal knowledge" of the theorems. And third, our model takes into account the fact that people do not infer all the facts they can infer. Immediate deduction stops as soon as the subject thinks that he has enough facts to handle the problem.

Summary

This paper reports an experiment on learning by problem-solving. During ten individual one-hour sessions, three third-grade secondary French school students

(12-13 years old) were asked to verbally demonstrate Geometry problems. A computer model was built that simulates the recorded verbal protocols. It was postulated that the students' knowledge can be analyzed into three components: strategic knowledge for using deduction rules in the process of proof generation; linguistic knowledge for interpreting the problem statements, the questions and the theorems stated in natural language; and intuitive knowledge in Geometry that is used in the processes of linguistic interpretation and proof generation.

References

Anderson, J.R. (1983). Acquisition of proof skills in Geometry. In R.S. Michalski, J.G. Carbonell, T. Mitchell (Eds.), Machine Learning: an Artificial Intelligence Approach (pp. 191-219). Palo Alto, Ca: Tioga.

Greeno, J.G. (1979). Theory of constructions and set in problem solving.Learrning Research and Development Center, Technical Report no.4, University of Pittsburgh.

Greeno, J.G. (1980). Instruction for skill and understanding in mathematical problem solving. Proceedings of the XXII International Congress of Psychology, Leipzig, July 1980.

Holland, G. (1983). Working backwards in solving geometric calculation problems. Proceedings of the Seven International Conference for the Psychology of Mathematics Education , Rehovot, Israel,1983.

Holland, G., & Reitz, H. (1985). Computer modelling of mathematical problem-solving process as an instrument of competence analysis. Proceedings of the Nineth International Conference for the Psychology of Mathematics Education , The Hague,1985.

Larkin, J.H., Mc Dermott, J., Simon, D.P., & Simon, H.A. (1980). Models of competence in solving Physics problems. Cognitive Science, 4, 317-345.

Nguyen-Xuan, A., & Grumbach, A. (1985). A model of learning by solving problem with elementary reasoning abilities. In G. d'Ydewalle (Ed.), Cognition, Information Processing, and Motivation (pp. 445-457). Amsterdam: North-Holland.

Nguyen-Xuan, A., Boudarel, M., Colomb, J., Garcelon, C., & Ladeveze, A. (1987). L'apprentissage du raisonnement deductif en geometrie. In J. Colomb,& J.F. Richard (Eds.), Resolution de Problemes en Mathematique et Physique (pp. 89-144). Paris: I.N.R.P.

Nguyen-Xuan, A. (1990). Deductive reasoning in geometry: search strategies and surface features. In H. Mandl, E. De Corte, N.S. Bennett, H. Friedrich (eds), Learning and Instruction: European research. Oxford, Pergammon Press.

HOW DO PUPILS LEARN MATH CONCEPTS?

Paola Forcheri, & Maria Teresa Molfino, Istituto per la Matematica Applicata del Consiglio Nazionale delle Ricerche, Genova, Italy

Knowledge based systems, able to assist students during the solution of cognitive problems, seem to be useful tools to make pupils formalize specific knowledge and develop logic-deductive reasoning at the outset of schooling. A crucial aspect in building these kinds of systems is to adapt the knowledge representation and the reasoning to the pupils' mind, and to realize a dialogue with the user which facilitates the learning. With these ideas in mind we developed a knowledge based educational system, called TELLER GUIDE, apt to be used in the teaching/learning of arithmetic.

Several pedagogical questions have been addressed through the TELLER GUIDE.

The first regards the choice of a class of problems which is both stimulating and familiar to pupils and, in the same time, require them to handle fundamental mathematics topics. To satisfy this need, TELLER GUIDE has been centered on a cryptoarithmetic game. The game is as follows: to find out what digits must be substituted to the symbol in a symbolic relation, such as, for example, WE+WE=ALL, so that the relation itself can be interpreted as an arithmetic sum. Different symbols correspond to different digits. The relation is presented in a column.

The second question is related to the design and implementation of a reasoning mechanism which reflects, at some extent, the way pupils usually tackle the class of problems at hand. To this end, the solution mechanism on which TELLER GUIDE is based upon has been designed based on an experimental analysis of the behaviour of a sample of pupils.

The results of this analysis led us to implement a solution mechanism which integrates two different kinds of reasoning, which alternate in a non-deterministic way, depending on the situation at hand. The first one, called "domain oriented", is based on the idea of modelling the relation by means of facts and rules. Facts represent the constraints on the relation (for example, referring to the relation WE+WE=ALL, W, E, A, L represent different digits; neither W nor A represent the 0 value, as they represent the most significant digit of a number greater than 9, ...). Rules represent the arithmetical knowledge which allows reasoning about the symbols which pertain to the same column and about the mutual relation between columns (for example, given the relation WE+WE=ALL, from the unit column we can immediately deduce that L is an even number as it is the double of a number;

by analysing the hundred column, the 1 value is immediately deduced for the symbol A as it represents the carrier of the tens column). The second kind of reasoning, called "general problem solving", is used to hypothesize a correspondence between symbols and digits and to verify, by substituting digits to symbols in the given relation, if the arithmetical equality is satisfied. This "attempt" strategy is employed when no rule can be applied to the situation at hand.
The third question addressed through TELLER GUIDE regards the design and implementation of an assistance module which gives explanations and suggestions suitable to pupils' knowledge and attitude. To this end, the TELLER GUIDE carries out an "educational dialogue" based on the idea that the concept formation process, at the age at issue, is achieved mainly by reasoning by analogy and counter-examples. The analogies and the counter-examples should be adapted to the history of the specific pupil. Given a relation to be solved (i.e. a game to be played), the pupil is asked to perform a move. A move can be a deduction, or an attempt. The pupil indicates that he intends to perform a deduction by choosing the column on which to operate; an attempt is indicated by choosing a symbol. The solution mechanism is then activated. If the move proposed by the pupil is an attempt, and there exists a rule that can be applied in the situation at hand (that is, a deduction can be performed), the move is not allowed. The system suggests the pupil that a deduction can be performed and asks for a column. If no rule is found that can be applied to this column, the error is pointed out and the pupil is asked for another column.
Otherwise, if the move proposed proposed by the pupil is a deduction, and there exists a rule that can be applied to the chosen column, the TELLER GUIDE shows a series of one-column relations, which constitute examples of admissible and not admissible deductions from the column. Each non-admissible deduction represents an arithmetic misconception or it is a generalisation of the column. The pupil is required to choose one or more relations. The answer of the pupil is used to update the system hypotheses about the level of learning of the concepts involved in the reasoning. In turn, these hypotheses guides the choice of the analogies and counter-examples presented in the following steps of the dialogue. The TELLER GUIDE will be presented in this paper. This analysis is aimed at giving an operative contribution to the computational understanding of the processes underlying knowledge communication.

References

Forcheri P., & Molfino M.T., (1989,). The Teller Simulator: an automatic problem solver for educational use, in Lecture Notes in Computer Science 360, Maurer H. (Ed.), Springer-Verlag, pp.117-131

Forcheri P., & Molfino M.T., (1991). Designing a knowledge-based learning environment for arithmetic concepts, Computers and Education, Vol.16, n.2, pp. 143-151

STRATEGIES IN LEARNING TO USE A COMPUTER APPLICATION PROGRAM

H.H. Leemkuil, Centre for Applied Research on Education (OCTO), University of Twente, Enschede, The Netherlands

Introduction

A commonly used strategy in teaching students to use a programming language or a computer application program like a spreadsheet, a database management system or a word processor, is to put them behind a computer and give them a textbook with expository information, problems and tasks. In this situation the students have the opportunity to apply the newly acquired information directly on the computer. Several researchers (Hiltz and Kerr, 1986; Carroll and Rosson, 1987; Van Merriënboer and Krammer, 1987; Van Dijk, 1988; Weida, 1991) have noticed that students in this situation do not read the information very well but focus their attention on the problems. When solving the problems they don't plan their activities but they just start typing. This continues until something does not work correctly, in which case the student looks for help from the textbook, the instructor, a friend or the help key. The students relie heavily on a trial-and-error method. This doesn't seem to be an effective strategy. Carroll and Rosson (1987) state that in designing training material one has three options, attack this strategy, try to limit the effects of it or, capitalize on the strategy. In the experiment presented in this paper two of these options will be investigated.

The computer application program which was used in this study is DataStat (Databases and Statistics). DataStat is a program which consists of two components: a relational database management system (RDBMS) and a statistical package. The RDBMS is used to select data from databases. The query language which is used in this system is based on SQL (Structured Query Langauge), the international standard language for relational databases. The selected data can be transferred to the statistical package to analyze the data. This package offers several possibilities to represent frequency distributions and relations between variables.

In the study reported here two treatments were used. In the first treatment a textbook was used to present the information about the use of DataStat to the student. The information was divided in 5 chapters. Following after the expository information of each chapter, were some problems which the students had to solve by using DataStat. While reading the information in a chapter the students could use DataStat, although this was not stimulated. In the second treatment the same information was presented to the student by means of an instructional

computerprogram in stead of a textbook. While working with this computerprogram the students could not use DataStat. The problems which had to be solved by using DataStat, were presented in a reference book, in which also an overview of the content of the databases was given. In this treatment the student had to read the information in a chapter of the instructional computerprogram first, then leave this program and start up DataStat to solve the problems. The students in the first treatment group have the opportunity to use the strategy they adopt spontaneously, while the students in the second treatment group could not use this strategy.

Method

Subjects

Subjects were 66 fifth grade high school students (5 vwo) from two different schools (two classes in each school). Each class has a different teacher. All of the students in one of the classes of each school were in treatment one and all of the students in the other class were in treatment two. Each student had a computer at his disposal.

Procedure

The experiment consisted of eight sessions. In the first session a test to asses field(in)dependence (a version of the Group Embedded Figures Test) was administered and general information about DataStat and the instructional material was given. In the next six sessions students worked with the instructional material and DataStat. In the last session a posttest was administered. Because of practical reasons (availability of the computers) this was a paper and pencil test. The students could not use a computer while making the test. The students were also asked to indicate if they had any experience with the query language SQL and to write down their math grade.

Results

In table 1 the mean scores on the posttest are given. As can be seen students in the textbook treatment group have a mean score that is 20 points higher than the other treatment group. An analysis of variance reveals that this difference is significant, $F(1,64) = 12.19$, $p = 0.001$. Table 1 also shows that this difference is found in both schools. These effects are also significant, school 1: $F(1,36) = 9.06$, $p = 0.005$ and school 2: $F(1,26) = 5.49$, $p = 0.027$.

Table 1
Mean scores and standard deviations (in parentheses) on the posttest for both treatment groups overall, and for each treatment group in the each of the two schools.

	Textbook	Computer
Mean posttest score	103.59 (24.96)	83.71 (21.27)
Mean score school 1	110.44 (22.23)	91.33 (25.41)
Mean score school 2	95.33 (25.41)	70.77 (21.77)

Table 1 furthermore shows that the students in school 1 have mean scores that are higher than those of the students in school 2. The reason for this difference between schools may be that a part of the students in school 1 already had some experience with the query language SQL in previous years (not with DataStat), while in school 2 almost everybody did not have this experience. To test if SQL experience is an important factor, the mean posttest scores were computed for those students who had SQL experience (mean score 108.04, S.D. 19.16) and those who did not (83.18, S.D. 23.71), regardless of the treatment group in which they were. An analysis of variance shows that this difference is significant, $F(1,64) = 20.42$, $p = 0.000$. So, SQL experience is an important factor. Therefore, it is important to investigate if both treatment groups differ in SQL experience. In the textbook group 44% had SQL experience and in the computer group 38%. This is a minor difference. So it can be concluded that SQL experience does not account for the difference between treatment groups.
The two other individual difference variables which were measured also don't seem to cause the difference between the groups, because the group means of these two variables, mathematics grade and field(in)dependence, are comparable. Mathematics grade mean scores are: 6.13 (1.66) and 6.06 (1.06). And the mean field(in)dependence-scores are: 15.69 (7.92) and 14.76 (5.99). The correlation coefficients of these two variables and the posttest-score are low: $r = 0.19$ for the correlation between mathematics grade and posttest-score, and $r = 0.26$ for the correlation between field(in)dependence-score and the posttest-score.

Discussion
The significant difference on the posttest score between the two groups seems to be caused by the treatment, and not by factors like experience and ability or by school or teacher related factors. In which way do the two treatments differ? They differ in two ways. Firstly, the instructional text is delivered by means of a textbook in the first treatment group and by an instructional computerprogram in the second group. Secondly, in the first treatment group students have the opportunity to work with DataStat while they are working through the instructional text and in the

second treatment group students don't have this opportunity. It is hypothesized that the second factor is the most important one. From observations it is concluded that most of the students in the first treatment group actually use DataStat while they are working through the instructional material. And although they don't read the material very well, they are actively engaged in working with DataStat, while the students in the second treatment group, by force, are more passive. The students in the first treatment group have the opportunity to use the strategy they adopt spontaneously, while the students in the second treatment group could not use this strategy. In accordance with Carroll et al. (1988) it is hypothesized that the second treatment group performs worse because there is a mismatch between the instructional material and the learning strategy students adopt spontaneously in this kind of situations. Carroll et al. advocate a training design that capitalizes on manifest learning strategies.

References

Carroll, J.M., & Rosson, M.B. (1987). Paradox of the active user. In: J.M. Carroll (Ed.), Cognitive aspects of human-computer interaction. Cambridge, Massachusetts: The MIT press.

Carroll, J.M., Smith-Kerker, P.L., Ford, J.R., & Mazur-Rimetz, S.A. (1988). The minimal manual, Human-computer interaction, 3, 123-153.

Dijk, E.M.A.G. van (1988). Probleemoplossen met SQL (Problem solving using SQL). Verslag van een vooronderzoek. Enschede: Universiteit Twente, Memorandum INF-88-58.

Hiltz, S.R., & Kerr, E.B. (1986). Learning modes and subsequent use of computer-mediated communication systems. Human factors in computing systems. CHI '86 Conference proceedings. Association for computing machinery.

Merriënboer, J.J.G. van, & Krammer H.P.M. (1987). Instructional strategies and tactics for the design of introductory computer programming courses in high school. Instructional Science, 16, 251 - 285.

Weida, N.C. (1991). Teaching students to think of spreadsheets as mathematical models. Journal of education for business, 66, 5, 283 - 286.

EVALUATION OF A METHOD OF TEACHING SKILLS: THE FEUERSTEIN'S INSTRUMENTAL ENRICHMENT TECHNIQUE IN LOW-QUALIFIED ADULT'S PRE-TRAINING SESSIONS

Even Loarer, Michel Huteau, Institut National d'Etude du Travail et d'Orientation Professionnelle, Service de Recherches; Jacques Lautrey, Université René, Descartes-Paris V, Laboratoire de Psychologie différentielle, Paris, & Daniel Chartier, Association nationale pour la Formation Professionnelle des Adultes, Direction de l'Orientation, Département Etudes et Méthodes, Bagnolet, France

Introduction

The Instrumental Enrichment Program (PEI - Programme d'Enrichissement Instrumental) of Feuerstein *et al*, (1980) is a method of remedial cognitive mediation the aim of which is to generate a "structural modification" in the cognitive abilities of individuals. The authors propose that this modification is brought about by the use of exercises (the instruments) and requires only a minimum of knowledge on the part of the individuals. The *instruments* used are, as far as possible, devoid of all context. The method relies on the learning of sufficiently general cognitive and metacognitive strategies so as to be easily transferable to other areas of diverse contents.

Moreover, among the particular attributes aimed for by the method, special care is accorded to the control of the subject's own impulsiveness (*One minute! Let me think*) and confidence building through a growing feeling of competence in the learners.

The present study aimed at evaluating the effects of PEI, both on cognitive and conative levels, among adults with low qualification levels

Particular attention is equally accorded to examining the transferability of the acquired skills, for this aspect constitutes the central issue raised by the concept of remedial cognitive mediation, it is this which fundamentally distinguishes it from simple training for specific tasks.

Method

Subjects

The subjects were 220 adults who had left school at 16 years of age, without qualifications. Due to their low level, it was necessary for them to follow a pre-

training course in order to reach the selection criteria required for training as a qualified workers. This full-time course lasted 15 weeks. It consisted of an initiation to professional activities and bringing them up to level in French and Maths. The average age of the subjects was 25 and about 85% of them were male. The sessions consisted of fifteen or so trainees in A.F.P.A. training centres spread throughout most of France.

Procedure

The subjects were evaluated on three occasions: at the beginning of the course (pre-test), at the end of the course (post-test 1), then six months after having left the course (post-test 2).

The experimental group (EG) received, in addition to the usual training given in this type of course, 4 sessions of 1 hour 30 min to 2 hours of IEP per week thus totalling 100 hours in all.

In order to neutralise any bias of the observations due to the introduction of an experimental process among the trainees (Hawthorne effect) and any effects due to the particular involvement of the trainers, the control group (CG) followed, at the same pace as the EG, additional work sessions on computers (Computer Aided Training and the use of word processing and spreadsheet software) this was presented to them as being a new training method. The trainers of this group received special training in order to teach these techniques.

Material

The tests were identical for both the pre-test and post-test 1.

Variables taken into account concerned the cognitive domain, personality, and the ability to plan actions.

Cognitive domain

- inductive reasoning was evaluated with the aid of a group test requiring the discovery of the odd man out among five symbols,
- verbal aptitude was measured through four tests comprising a verbal analogy test, a vocabulary test, a phrase comprehension test and a verbal logic test,
- patial aptitude was evaluated using a cross section of volume test,
- creativity was tested by requiring the subject to produce a list of objects corresponding to a certain definition.

In the area of personality, we administered:

- a self-image questionnaire, a re-worked and adapted version of the Gordon personality Inventory,
- a locus of control questionnaire (an adaptation of the Rotter questionnaire),
- an individual test of Reflexive-Impulsive Style: the Kagan MFFT.

Planning

Two individual tests were specially elaborated for the experiment.

In the planning ability test of a workshop task the subject was asked to make a list of the tools, materials and operations required to produce a particular object (in wood or in sheet metal) that was shown to him. Note was taken of the time taken before beginning to draw up the lists. A coefficient for the quality of the response

was calculated by taking into account the conformity of the list produced with a standard list established by professionals.
In the delivery planning test, the subject was required to establish a delivery route for seven packages, taking into account a certain number of constraints (times of day, distances, etc.) Again, the initial time taken for reflection was noted as well as the quality of the answer in relation to the optimum itinerary.
At the post-test 2, only the cognitive tests were given, as well as a single simplified questionnaire dealing with self-image and locus of control.

Results

Effects observed at post-test 1

The results concern 178 subjects, present for both series of tests (92 of the EG and 86 of the CG). The averages of the groups were compared by analysis of covariance procedure, by taking the pre-test score as covariate.
The effect size was calculated by dividing the difference of group means by the mean of their standard deviations.
On the level of cognitive variables, a significant increase of EG progress is observed compared to the CG: reasoning ($p<0.01$), verbal aptitude ($p< .001$) and spatial aptitude ($p< .06$). The size of the effects are respectively .24, .28 and .20.
On the other hand, no difference is observed between the groups at post-test 1 for the creativity dimensions (fluency, flexibility and originality).
On the level of the personality,a difference is observed, at the advantage of EG ($p< .02$) in the development of the self-image: subjects in this group show, at the post-test, a higher sef-esteem. Regarding the internalising of control, there is no difference between the two groups.

Transfer of acquisitions to different tasks

On the individual Reflexive-Impulsive Style test, the EG subjects progressed more than the CG: at post-test 1 they took longer to reply ($p< .003$) and made fewer errors ($p<.001$). In the planning of a professional task, they devoted more time to reflection ($p<.02$) but produced replies whose quality was not greater than that of the CG. In the delivery-planning test the two groups showed no difference in time taken for reflection or in the quality of the route produced.

Effects observed at post-test 2

The results were calculated on a limited sample of subjects having undergone all three series of tests (50 for the EG and 40 for the CG).
For these subjects, the progress observed at post-test 1 was maintained without the performances increasing: there was no significant progress between post-test 1 and post-test 2, nor any significant differences between the EG and the CG at post-test 2.

Discussion

The results suggest that the differential effects of IEP are essentially manifested in an significant improvement in classic aptitude scores. The differences are, however, of moderate amplitude since the observed effects imply an overlap between score distribution of the EG and CG in the order of 30 to 40%.

The acquisitions of both groups were maintained six months after the course, but the long term growth, claimed by the authors of IEP, has not been proved.
The fact that the progress made in the impulsiveness - reflection and intelligence tests were not transferred to the delivery planning task, which benefited from no training, leads one to doubt that the general strategies of problem solving have been really acquired, and brings into question the structural character of the cognitive modifications offered by the method.

CONCEPTUAL GROWTH AND CHANGE THROUGH ANALOGICAL REASONING AND METACONCEPTUAL AWARENESS

Lucia Mason, Department of Educational Sciences, University of Padova, Italy

Abstract

This study is an investigation on the role of analogy in learning and restructuring science concepts, and on the relation among analogical reasoning, metaconceptual awareness, and conceptual change. The analogy introduced to children of elementary school, divided into two experimental conditions, puts the new concept of clorophyllian photosynthesis in familiar terms. Quantitative, but mostly qualitative data analysis is presented in order to underline some aspects of anomaly creation in current mental models and of change of explanatory frameworks, as well as of misrepresentations of the target domain induced by the analogical mapping. These aspects are related to children's productive use of analogy and metaconceptual awareness about their knowledge construction activity. Some implications from the standpoint of instructional design are introduced.

Theoretical Framework

Meaningful learning is considered a generative process in which students construct relations among and within concepts (Wittrock, 1985). The mapping of knowledge from a familiar domain (the source) to a less-known one (the target), that is analogy, is a powerful tool in thinking and learning concepts relationally. In this way, organized networks of related information, rather than lists of unrelated facts, are acquired. Many theoretical studies have pointed out the strong implications of analogies as powerful tools in teaching and learning. They facilitate: the coding of new knowledge; the acquisition and organization of information; the access to and the retrieval of knowledge previously stored in memory; the creation of anomalies in a conceptual network or new schemata; the restructuring of a system of concepts involved in a mental model, by overcoming misconceptions through the modification of entrenched beliefs (Brown & Clement, 1989; Rumelhart & Ortony, 1978; Stepich & Newby, 1988; Vosniadou & Ortony, 1990).

However, the process of conceptual change (Carey, 1985) seems to require metaconceptual awareness in students, in order for them to be conscious of their own deep beliefs which must be considered questionable, and which may result

inadequate as hypotheses about reality (Carey et al. 1989; Kuhn, Amsel, & O'Loughlin, 1988; Sodian, Zaitchik, & Carey, 1991).
In this work two questions are addressed: a) what are the role and the effects of analogy in children's processes of conceptual growth and restructuring about a science topic?; b) is there any relation between metaconceptual awareness and productive use of analogy in those processes?
Some interesting experimental researches have investigated the use of analogies in teaching and learning concepts, but almost all the studies were carried out in simplified contexts (laboratory conditions), without taking into account the complexity of a natural learning environment such as a classroom. I think it is very important to focus on the teaching-learning by analogy in natural and relevant conditions, that is an ordinary school setting. Moreover, it is useful to face complex concepts held in curricular disciplines, to improve the ecological validity of the research itself.

The Research

Aims. In the light of data presented in two our previous studies, showing that analogy is an effective learning aid, the purposes of this research are the following:

- confront the effects of analogical comparisons drawn by pupils themselves and analogical comparisons clearly explained by the experimenter, in constructing and restructuring science knowledge;
- evaluate the impact of pupils' knowledge structures on the productive use of analogy;
- investigate the metaconceptual awareness in pupils by giving them the opportunity of reflecting on the process of elaborating knowledge in a curricular field (the biological one) while learning facts and principles. The specific goal is to provide children with situations in which they can realize that their beliefs are constructions subject to falsification, and not "true facts";
- to explore the relation among pupils' metaconceptual awareness and use of analogy in the processes of conceptual growth and restructuring.

Method

Subjects

The sample is made up of 44 children (aged between 10 and 11 years), attending their last year at elementary schools in Padova, divided into two experimental groups.

Material

The specific material consists in the analogy between making a cake and the clorophyllian photosynthesis (Glynn, 1990) aimed at teaching that plants are the only living organisms able to produce by themselves the food they need (autotrophism), by changing raw materials into end products (sugar and oxigen) so as to carry out their body functions and to grow. Plants release oxygen in the air as a waste product, so they are absolutely necessary for other living beings. This is fundamental for the pupils to understand why no plant must be destroyed. A good example is the large support given by many people of various nations to the save-the-Amazonian forest campaign.

Some scientific experiments about living functions in plants and about clorophyll are performed in the classroom before the analogy is introduced. The main functions of this analogy, as effective learning aid, are three: concretizing, active assimilation and structurizing (Simons, 1984). Concretizing function makes abstract information more concrete and imaginable. Structuring function is related to the role of an existent structure which can form the basis for a new schema. Active assimilation function facilitates students to integrate the new and previously constructed knowledge in the conceptual structures.

Procedure

The children are individually interviewed on some questions (factual, explanatory and generative) on the conceptual area to be investigated in order to ascertain their prior knowledge (Stavy, 1987; Wandersee, 1983) and, moreover, on their metaconceptual understanding about the nature and construction of their own beliefs and knowledge. Then the experimenter organizes the usual school setting for four sessions (previously agreed with each class teacher) in which the biological topic is dealt with. The analogy is introduced at a certain point of the didactic intervention, under two different conditions:

- in the first one all the analogical mapping between making a cake and clorophyllian photosynthesis is clearly given to the children;
- in the second one the children are stimulated to detect all the relations between the two processes, after they have been told that an analogy can be drawn.

The pupils are engaged in a process of analogical reasoning in a learning environment in which large and small group discussions are encouraged.
At the end of the intervention in the classroom, and again twenty days later, each child's conceptual structures are evaluated through different tasks: short text, picture, explanation questions, transfer questions, multiple choice questionnaire. Moreover, he/she is interviewed again on the same set of questions presented at the beginning, and on other questions related to his/her understanding and use of the analogy in the conceptualization process.

Data analysis

For the quantitative analysis, the information contained in the various tasks was scored on the basis of its appropriateness and complexity. The qualitative analysis was based on the transcription of the structured interviews and discussions audio-taped in the classroom.

Results

The results can point out that the context as well the way of elaborating an analogy play a role. Moreover, the level of metaconceptual awareness is related to both the productive use of the analogy and the change of interpretative frameworks to process new information and to resolve conceptual anomalies. The qualitative analysis is of particular interest and rich of educational implications.

Discussion

The study shows the influence of entrenched beliefs in the knowledge acquisition process. In order to change the pupils' existing conceptual structures, it is not enough to present them with new information as a given fact, since they need to discover, or receive, a new explanatory framework. In some contexts, analogy can be a powerful tool in generating anomalies within a current mental model or in creating a new one. At the same time, analogy may induce in the pupils some misrepresentations of the target domain.

Furthermore, the analogical reasoning itself, as a fundamental way of constructing scientific knowledge, can be fostered by a kind of teaching aimed at enhancing metaconceptual learning.

References

Brown D.E., & Clement J. (1989). Overcoming misconceptions via analogical reasoning: Abstract transfer versus explanatory model construction. Instructional Science, 18, 237-261.

Carey S. (1985). Conceptual change in childhood. Cambridge, Ma.: MIT Press.

Carey S. et al. (1989). "An experiment is when you try it and see if it works": A study of grade 7 students' understanding of the construction of scientific knowledge. International Journal of Science Education, 11, (5), 514-529.

Glynn, S.M. (1990). The teaching-with analogies (TWA) model: Explaining concepts in expository texts, in K.D. Muth (Ed.). Children's comprehension of text: Research into practice. (pp. 185-204). Newark, De.: International Reading Association.

Kuhn D., Amsel E., & O'Loughlin M. (1988). The development of scientific thinking skills. New York: Academic Press.

Rumelhart D.E., & Norman A.D. (1978). Accretion, tuning and restructuring: Three modes of learning. In J.W. Cotton & R.L. Klatzky (Eds.), Semantic factors in cognition (pp. 37-53). Hillsdale, N.J.: Erlbaum.

Simons, P.R.J. (1984). Instructing with analogies. Journal of Educational Psychology, 76, (3), 513-527.

Sodian B., Zaitchik D., & Carey S. (1991). Young children's differentiation of hypothetical beliefs from evidence. Child Development, 62, 753-766.

Stavy R. (1987). How students aged 13-15 understand photosynthesis. International Journal of Science Education, 9, (1), 105-115.

Stepich D.A., & Newby T.J. (1988). Analogical instruction within the information processing paradigm: Effective means to facilitate learning. Instructional Science,17, (2), 129-144.

Vosniadou, S., & Ortony A. (Eds.) (1989). Similarity and analogical reasoning. Cambridge, MA.: Cambridge University Press.

Wandersee J.H. (1983). Student's misconceptions about photosynthesis: A cross-age study. In H. Helm & J.D. Novak (Eds.). Proceeding of the international seminar: Misconceptions in science and mathematics. (pp. 459-484). Ithaca, N.Y.: Cornell University Press.

Wittrock M. (1985). Learning science by generating new conceptions from old ideas. In L. West & A. Pines (Eds.), Cognitive structure and conceptual change. Orlando. Fl: Academic Press.

GOALS AND EFFECTS OF LEARNING STRATEGIES COURSES

Jan Kaldeway, & Fred Korthagen, IVLOS, Rijksuniversiteit Utrecht, Utrecht, The Netherlands

Introduction

This study is part of a research programme aimed at the evaluation of learning strategies courses. In the first stage of this program a model for categorising study activities was developed (see Kaldeway & Oost, 1990, 1991) based on current theory on metacognition (see Vermunt, 1989, for a review). In the study reported here, which is the second stage of the programme, the focus is on evaluation studies of learning strategies courses in tertiary education, more specifically the relationships between goals and effects of these courses. On the basis of the results of this study further research will be done on the evaluation of a learning strategies course of the University of Utrecht.

In several reviews it is concluded that learning strategies courses do have effects (e.g. Kulik et al, 1983; Weinstein & Mayer, 1986; Pintrich et al, 1986). However, a review of the relationships between the goals and effects of learning strategies courses doesn't exist. Insight into these relationships is necessary in view of designing learning strategies courses.

The central question of this review is: what are desirable goals of learning strategies courses, i.e. goals that contribute most to the study results of the students? Are these goals the fostering of planning or concentration, the development of adequate study methods, enhancing self–esteem or are other important goals involved? In order to be able to answer this question we have to know what factors have more effect on study results than others, and, if some factors (such as concentration and self–esteem) may be influenced by others (such as study methods).

Method

The selection of research reports on learning strategies courses was based on an *Eric search*, the *British Education Index*, on a number of relevant scientific journals, from 1986 to 1991, and on reviews and conference proceedings on learning strategies (Pintrich et al, 1986; Weinstein & Mayer, 1986; Kirby, 1984; Weinstein et al, 1988; Schmeck, 1988). Courses were selected on two criteria. First, they had to be in tertiary education, and second, their effect should have been assessed. In this way, sixteen courses were selected.

To analyse the goals of the courses, a classification was used based on research on educational achievement factors in tertiary education. The factors study method, self–esteem, study conceptions and study motives were identified as possibly important *determinants of study success.*
The factor *study method* was specified in four categories of primary study activities and two supporting study activities. For the analysis of primary study activities the model developed by Kaldeway & Oost (1990, 1991) was used. In this model the so–called metacognitive activities are specified and interrelated. The model distinguishes four categories of study activities, *strategic study activities* (formulating the general study goals), *operational study activities* (designing a study plan), *executive study activities* and *reflective study activities* (controlling the study process and the study results).
The supporting study activities are *concentration management* and *time management.*
To analyse the effects of the courses, the same classification was used. Three effect measures were added: reading comprehension, examination results and time investment.

Results
The most important findings are the following (see the appendix for an overview of the results). The most striking finding is the fact that learning strategies courses mainly aim at *executive study activities*, such as schematising, networking or annotation. Nine out of the sixteen courses analysed were exclusively focused on executive study activities or on executive and reflective study activities. Strategic study activities formed part of only one course, operational study activities formed part of five courses.
As for the *effects* of the courses it is notable that significant effects in reading comprehension, examination results or time investment are reported on almost all courses.

Conclusion
On the basis of the analysis of the courses some general conclusions can be drawn.
First, significant effects of the courses appear irrespective of the specific goal of a course. This applies to limited courses focused on one specific executive study activity, such as networking, as well as to wide–ranging courses directed to operational and executive study activities, concentration management, time management, self–esteem and study conceptions.
Secondly, only limited research has been done on the effects of the courses on *study results.* Most evaluation studies measure effects on reading comprehension but do not on examination results or time investment.
Thirdly, the conclusion can be drawn that on the basis of existing research on learning strategies courses no final statements are possible with regard to the relative importance of the different goals. Neither are statements possible with regard to the interrelationships between these goals, such as the influence of training in study methods on concentration or self–esteem.

Discussion

On the basis of the available research on learning strategies courses it is not possible to formulate specific directions for designing a learning strategies course. However, it is possible to formulate more specific research questions aiming at gaining more insight into the relationships between goals and effects of these courses.

First, research is needed on the relationship between the goals of learning strategies courses and the effects on *study results* (examination results and time investment). It is possible that existing courses which have effects on reading comprehension, do not have effects on study results because no attention is paid to strategic and operational study activities.

Secondly, the interrelationship between the different goals of learning strategies courses have to be investigated more precisely. For example, what are the effects of training in study activities on concentration, self–esteem, study motives and study conceptions? In addition, in this research the role of individual differences between students has to be investigated, as well as differences in type of institution (e.g., college or university) and in discipline (e.g., science or literature).

References

Bronders, M. (1991). Studiemethoden: een onderzoek bij eerstejaarsstudenten – de opbouw van een schematiseertraining in IMG. Tijdschrift voor Hoger Onderwijs, 9 (2), 84–93.

Dendato, K.M., & Diener, D. (1986). Effectiveness of cognitive/relaxation therapy and studyskills training in reducing self–reported anxiety and improving the academic performance of test–anxious students. Journal of Counseling Psychology, 33 (2), 131–35.

Jacobowitz, T. (1990). A metacognitive strategy for constructing the main idea of text. Journal of Reading, 33 (8), 620–624.

Kaldeway, J., & Oost, H. (1990). Een proces–produktmodel van het studeren. In: F.J.Slobbe, J.Blankestijn, O.S.C.Zeeuwen & T.J.M.Willemsen (red.), Terugblik op 10 jaar studievaardigheden: toekomstige ontwikkelingen (pp.153–175). Proceedings Landelijke Dag Studievaardigheden.

Kaldeway, J., & Oost, H. (1991). Strategisch studeren kun je leren. *Onderzoek van onderwijs, 20 (2)*, 28–30.

Kirby, J.R. (1988). Style, Strategy, and Skill in Reading. In: Ronald R.Schmeck (ed.), Learning Strategies and Learning Styles (pp.229–273). New York: Academic Press.

Kulik, C–L.C., Kulik, J.A., & Shwalb, B.J. (1983). College programs for high–risk and disadvantaged students: a meta–analysis of findings, Review of Educational Research,53 (3), 397–414.

Lundeberg, M.A., (1987). Metacognitive aspects of reading comprehension: Studying understanding in legal case analysis. Reading Research Quarterly. 22 (4), 407–432.

Morgan, M. (1981). Self–derived objectives in private study. Journal of Educational Research, 74 (5), 327–332.

Nist, S.L. ,& Simpson, M.L. (1989). PLAE, a validated study strategy. Journal of Reading ,33, (3), 182–186.

Pintrich, P.R., Cross, D.R., Kozma, R.B., & McKeachie, W.J. (1986). Instructional Psychology. Ann. Rev. Psych., 37, 611–651.

Dansereau, D.F. (1985), Learning strategy research. In: Segal, J., Chipman, S. & Glaser, R. (eds.), Thinking and learning skills: relating instruction to basic research, 1:pp.209–240. Hillsdale, NJ: Erlbaum.

Schmeck, R.R. (ed.) (1988). Learning strategies and learning styles. New York: Plenum Press.

Simpson, M.L., & Nist, S.L. (1990). Textbook annotation: An effective and efficient study strategy for college students. Journal of Reading, 34 (2), 122–129.

Van Overwalle, F. (1989b). The effects of attribution–based intervention and study strategy training in academic achievement in college freshmen. Paper presented at the Onderwijs research Dagen, Leiden, June 4.

Vermunt, J.D.H.M. (1989). The interplay between internal and external regulation of learning, and the design of process–oriented instruction. Paper presented at the third Conference of the European Association of Research on Learning and Instruction, Madrid, September 4–7.

Weinstein, C.E. (1988). Assessment and training of student learning strategies. In: Ronald R.Schmeck (ed.), Learning Strategies and Learning Styles (pp.291–316). New York: Plenum Press.

Weinstein, C.E., & Mayer, R.E. (1986). The teaching of learning strategies. In: M.C.Wittrock (ed.), Handbook of research on teaching (pp.315–327). London: Macmillan.

Appenix - Goals and effects of learning strategies courses

	Goals										Effects											
	S	O	E	R	C	T	Z	M	V		S	O	E	R	C	T	Z	M	V	P	L	I
Goldfried et al 1978							x										x					
Dansereau et al 1979		x	x	x	x	x							x		x		x		x	x		
Holley et al			x																	x		
Collins et al 1981							x													x		
Brooks et al 1981			x																	x		
Morgan 1981	x	x																		x		
Cook 1982			x																	x		
Dansereau 1985			x																	x		
Dendato & Diener (*) 1986			x x				x x										x x o				x o o	
Lundeberg 1987			x																	x		
Weinstein 1988		x	x		x	x	x		x				x				x		x		x	
van Overwalle (*) 1989		x x	x x	x x			x x														x o o	
Nist & Simpson 1990		x	x	x																	x	
Simpson & Nist 1990			x	x																x		x
Jacobowitz 1990			x	x																x		
Bronders 1991			x										x							o		

S = strategic study activities
O = operational study activities
E = executive study activities
R = reflective study activities
C = concetration management
T = time management
Z = self-esteem
M = study motives
V = study conception

P = reading comprehension
L = exam results
I = time investment

(*) three experimental conditions
x goals: the course focuses on this factor
effetcts: significant effects on this factor were measured
o effects were measured, but no significant effects were found

EFFECTS OF SELF SELECTED STUDY STRATEGIES ON FIFTH GRADE STUDENTS' RECALL, READING COMPREHENSION AND DURATION OF STUDY

A. Munire Erden, & Melek Demirel, Hacettepe University, Faculty of Education, Ankara, Turkey

Introduction

In the learning teaching process, what the students do to learn is as important as the quality of instruction provided. Each student has his own strategy in learning a new task. The strategy that the students use affects not only the process of selecting and organizing the new knowledge and integrating it into the memory, but also the learning outcome and the quality of performance.(Weinstein E.C. and Mayer E.R.,1986)

A lot of research has been done on the effectiveness of the study strategies. Fowler and Barker (1974) and Rickards and August (1975) showed that studying a text by underlining it was more effective than mere reading. Annis (1985) found that the students who summarized the text in their own words were more successful on the test (which is based on the text) than the ones who did mere reading. Dyer, Riley and Yakovich (1989) compared note–taking, summarization and re–reading strategies and stated that these strategies had equal effects on learning.Most of the research on study strategies has been carried out on high school or university students. Besides, in these research generally the experimental method has been used and the students in the experiment groups have been asked to use a certain strategy. But learning strategies are acquired, developed and/or improved through their experiences over a long period of time.Brown and Smiley (1978) indicated that the students who used their own strategies(without having been directed) were more successful the most important information from a given text as compared to the ones who were asked to use a certain study strategy by the researcher.

The aim of this study is to investigate the effects of self–selected study strategies on recall, reading comprehension and duration of study of fifth grade students.

Method

The research was conducted on 75 students selected at randomly among the fifth grade students attending two primary schools in Ankara. In the research, a text, a related recall test, a reading comprehension test and an observation schedule were used as instruments to collect data. A text of approximately 500 words on the characteristics and flying principles of balloons was prepared. The recall test on

this text involved 34 items and its KR 21 reliability coefficient was found to be .75. The reading comprehension test involved 30 items and its KR 20 reliability coefficient was found to be .84.

The subjects were taken into the study room in groups of five and were asked to study the text as they wished and to use the blank study sheets in case they needed. The students wereobserved while they were studying and their duration of study and concentrations on the text were recorded. The recall test was given immediately after they finished studying. All the students took the reading comprehension test 3–4 days later. In addition, the study sheets and observation schedules were examined and categorized according to the study strategies observed on them.

Results

In analyzing study materials and observation schedules, five study strategies were found. The strategy most frequently used was mere reading (%41.3) followed by summarizing the text material (%33.4), underlining (%9.3), organizing (%9.3), and copying(%6.7).

In analyzing the data, one–way analysis of variance was used in order to determine the effects of study methods on recall, reading comprehension and duration of study. To find the source of the invariability, Scheffe test was used.

Results of the Analysis of Variance

- Recall: The result of the one–way analysis of variance on the data for recall test yielded a significant difference for study strategies (F:4.91, p <.01). The most effective strategies for the students at this level were organizing and underlining. However, no significant difference was found between the means of the test scores of students using these two strategies. The group with the lowest mean was the group who used copying as a strategy.
- Reading Comprehension: The result of the one–way analysis of variance on the data for the reading comprehension test yielded a significant difference for study strategies (F:3.89, p <.01). Mean scores of the students who used organizing and underlining strategies were found to be significantly higher than those of the students who used the other strategies. The detailed analysis showed that the students with poor reading comprehension didn't use the underlining and organizing strategies. Among these students, those who used summarizing and rereading strategies were found to be more successful as compared to those who used copying as a strategy. The most effective strategies for students with good reading comprehension were underlining and organizing.
- Durations of Study: The result of one–way analysis of variance on the data for the duration of study yielded a significant difference for study strategies (F:2.54, p <.05). The strategy which required the shortest duration was mere reading. The difference between the average durations of underlining and organizing strategies was not statistically significant. The longest duration required was that of copying strategy which was the least effective one.

Conclusion and Discussion

The study showed that the use of various study strategies by the students have different effects on recall, reading comprehension and duration of study. It is concluded that the most effective strategies for students at this level are organizing and underlining. However, since the students capable of using these strategies were very few and since mere readingstrategy was the most frequently used one, we can conclude that the students are unaware of the effective study strategies. Because the use of these study strategies is an important factor which affects achievement and decreases the duration of time devoted to study, these strategies must be taught to the students. The researches which were done up to the present time shows that organization and summarizing strategies can be taught. Pressley, Johnson and others (1989)indicated that the groups which were taught the summarizing strategy had better achievement than the control group.However, some researches indicates that the study strategy instruction over a short period of time is not effective (Bretzing, Kulhavy and Catering , l987;Thornton and others, 1986). Based on the results of these investigations, it can be concluded that the study strategies can be taught to the learners but it is clear that this learning process takes a long period of time. For this reason, from the first years of primary education on serious efforts to teach and develop study strategies are required. In order to be successful in this, the teachers must certainly be informed about these strategies and trained on the instruction of them.

References

Annis, L. F. (1985). Student–generated paragraph summaries and the information processing theory of prose learning.Journal of Experimental Education , 54, 4–10.

Bretzing, H. B., Kulhavy W. R., & Caterino, C. L. (1987). Notetaking by Junior High Students.Journal of Educational Research, 359–362.

Brown, A. L., & Smily, S. S. (1978). The development of strategies for studying texts. Child Development, 49, 1076– 1088.

Dayer, J.W., Riley, J., & Yekovich F. R. (1989). An analysis of three study skills: Notetaking, summarizing and rereading. Journal of Educational Psychology, 73, 3–7.

Fowler, L. R., & Barker, S. A. (1974). Effectiveness of highlighting for retention of text material.Journal of Applied Psychology,. 59, 358–364.

Rickards, J., & August, G. J. (1975). Generative underlining strategies in prose recall. Journal of Educational Psychology, 67, 860–865.

Pressley, M., Johnson, C. J., Symons, S., Mcgoldrick, J. A., & Kurita, J. A. (1989). Strategies that improve children's memory and comprehension of text. The Elementary School Journal, 90, 3–32.

Thornton, E.N. , Bohlmeyer M. E., Dickson A. L., & Kulhavy W. R. (1986). Spontaneous and imposed study tactics in learning prose. Journal of Experimental Education, 111–124.

Weinstain, C. E., & Mayer, R. E. (1986). The Teaching of learning strategies (p. 315–327). in M. C. Wittrock (ed.), Handbook of research on teaching, New York: Macmillan.

SPONTANEOUS AND CONSTRAINED STRATEGY USE IN LEARNING TEXTS

Ewa Czerniawska, Faculty of Psychology, University of Warsaw, Poland

Introduction

Many investigations,in the previous years,were focused on finding the best strategies for different memory tasks and constructing,on this basis,intervention programs. But a great deal of studies demonstrated that it was not easy to succeed in training strategies. Even if the short-term results were good, most often the transfer to real-life situations was not observed. The reason for that failure may be of various kind, for example, factors linked with emotions, motivation, mind-orientation, personality pointed out in Boekaerts', Schmeck's, Simon's or Wong's works, and cognitive determinants related to cognitive style or metamemory. Investigations in the area of metamemory show that children are gradually acquiring knowledge about determinants of memory effects and the ability to adapt their strategic behavior to the learning goals (cf.Brown, Flavell, Johnson). It can be expected that older children will be more likely to employ strategies better appropriate to the task. Nevertheless this is not always the case.
Some important questions have to be answered: do the objectively best strategies really lead to higher results, especially in the case of subjects presenting a definite learning style? What is preferable: the training of "better" strategies or the arousal of self-awareness and self-acceptation of one's own individual learning style?

Aim

The main aim of the research, presented below, was to supplement existing knowledge on memory strategies, by comparing the effects of spontaneous and constrained strategy use in learning texts by pupils from 5th to 8th grades, differing in declared "strategicity".

Subjects

Subjects were 168 pupils from Warsaw primary schools (5, 6, 7 and 8 graders), 21 high-strategic (HS) and 21 low-strategic (LS) at each grade level. They were selected on the basis of a questionnaire concerning strategies of learning texts. As HS were classified pupils who declared employing different and active strategies in the case of various school subjects (biology, geography and history); as LS-those who declared employing the same strategy independently of the school subject and their strategies were more surface.

Method

In the course of individual sessions, subjects learnt short texts (about 160 words each) under four conditions differing in strategy source (spontaneous choice versus constrained use) and in presented learning goals (reproductive versus productive). Immediately after learning each text pupils were asked two questions concerning the text: one reproductive and one productive.

Results

The analysis of strategic behavior showed differences depending on "strategicity" and grade level. LS subjects from grades 5, 6 and 7, less frequently than HS ones, modified their strategies in response to the presented learning goals and the chosen strategies were mainly low active. In the 8 graders the majority of subjects did not modify their strategies in the case of different learning goals but the employed strategies were more active and complex.

Multiple variance analysis indicated that memory effects were mainly determined by grade, "strategicity" and memory task, but not by their interaction. The pattern of influence was as follows:

- the results of 5 and 6 graders were significantly lower than those of 7 and 8 graders;
- HS subjects achieved significantly better than LS ones in all tasks;
- the overall reproductive effects were significantly higher in the "spontaneous strategy choice" situation than in the "constrained" condition; productive effects were also higher but not at a statistical level;
- subjects achieved better when the learning goal and the task measuring memory effects were concordant (reproductive-reproductive or productive-productive);
- in the "constrained" conditions the results were higher in the reproductive learning goal situation, than in the productive learning goal one.

Discussion

The obtained results showed that the possibility of spontaneous strategy choice, when different learning goals were presented, did not lead to the modification of strategic behavior in case of all pupils. "Strategicity", or learning style, and age seemed to be the most important factors determining the strategy choice. Younger pupils were more likely to change strategies if they were HS, older HS changed their strategies less often but more HS and LS subjects at this age level employ active strategies.

The strategies proposed in the "constrained" conditions were well appropriate to the learning goal: "repeating" to the reproductive goal ("remember as much as you can") and "making plans" to the productive goal ("make inferences on the basis of the text"). Previous investigations on learning texts showed that those strategies were well known and often employed by pupils in school. Also laboratory researches confirmed their effectiveness in different memory tasks.

My research indicates that a less active strategy is favorable for the reproductive effects but also shows that proposing an active strategy may disturb reproductive and productive effects. When active strategies are the effects of a spontaneous choice they lead to higher both reproductive and productive results.

Conclusions

- The research indicated that the following factors were main determinants of learning effects: age, "strategicity", strategy source and the relation between learning goals and memory tasks.
- pontaneous strategy choice led to higher results even if the chosen strategies were not "objectively" better appropriate to the task, especially for older and high-strategic subjects.
- When training strategic skills we must pay special attention to the pupils' learning style, trying rather to arouse self-awareness than propose "objectively" better strategic behavior.

A PROPOSITION OF EVALUATING A CHILD'S DEVELOPMENT IN THE PROCESS OF PRIMARY EDUCATION TRAINING

M. Jakowicka, Pedagogical University, Poland

Introduction

The question under consideration is a thesis that education means creating conditions for a child's development, organizing its activity on a higher and higher level of independence in situations which release the child's thirst for knowledge, curiosity and exploration. These situations being a source of the discoveries contribute to the child's cognitive, social and emotional development.

The following questions need to be asked: what to do, what kind of strategies and detailed practical solutions can be used so as to in the teaching-learning process enrich different traits of personality and to recognize and be able to function in real life in a happy, internally motivated way?

One of the important issues both for theory and practice are the indicators for evaluating the child d development. They seem necessary for the teacher in his process of instruction. The indicators should be perceptible as on their basis the child's progress is observed and evaluated. They should also integrate the comprehensive process of education not only the particular tasks of each school subject as it happens at present.

In the light of the idea "education to help development" we need to analyse the change occurring within the pupils by taking into consideration two levels of generalization. The lower "subject-oriented" one is supported by research and practical material. The higher level of "development" however, needs to be elaborated. A hypothesis can be put forward (it needs verification) that if a pupil makes progress in a general sense then he acquires skills in each of the school subjects (apart from the partially disordered cases).

Two important questions become a consequence of the presented problem:

- what kind of externally visible scopes of a child s behaviour can illustrate the progress in development?, what kind of scales and indicators can be applied;
- which of the analytical methods would be of value to be put in practice in the school conditions by the teacher and which, on the other hand would not cause disturbance in the process of teaching?

Three scopes can be applied here:

- the characteristic features of a child s activity during task performance;
- features of works of art created by the pupils;

- the classificatory functions which being applied to the process of education help to integrate the contents and ways of acquiring knowledge.

Each of the mentioned points refers to the work done on the contents of different school subjects which at the point of integration reach the level of activity which somehow stands apart from the detailed contents of knowledge.
The suggested conception of solving the problem was verified in a practical and theoretical way by a group of people working at the Pedagogical University in Zielona Gora. The results seem to confirm the right direction of research. The material was gathered from a test group which consisted of 700, nine-year-old pupils (i.e. the lower level of primary education) coming from different backgrounds. The material was compiled through observation and analysis of children's works of art. It was conducted in natural conditions of the process of education.
The paper will also include the analytical conception of a child's development, the results of the research in a global way and differentiated according to contents and environment both treated as independent variables.

Methods of analysis of a child's development in the process of education
The useful methods in the light of the established conception are: the observation of the child's activity during task solving and the analysis of the child s art creations. The main value of the method appears to be the application of it during the educational process without disturbing it. It enables us to watch the work ability, the overcoming of the difficulties and the change of behaviour at the higher levels of education. Similarly,the works of art are a result of integrated psychological, motivated and volitional activities. The methods, therefore, enable us to analyse everything which is comprehensive which means development.

The proposition of evaluating the development of pupils
Analysis of children's activity
The activity, its rationality and efficiency is one of the conditions to evaluate a man in different spheres of life. In literature there is an agreement as to the essence of the concept which means a conscious and intentional sequence of activities in order to fulfil a task.
The evaluation of the level of such a complex activity can be accomplished through analysis of specific features in different tasks. The dominant behaviour is taken into the consideration here. In the research the pupils solved 10 problems in four subjects; their native language, mathematics, art and natural science.
The proposition includes indicators and scales for evaluating the level of readiness, persistence and the speed considered as activity features and correctness considered as work feature.
The results indicate a high level of readiness and persistence,show a lower level of correctness and the lowest of all the speed level.

The analysis of the works of art
In the creation of a work of art an individual approach is present.
That is the taking advantage of accumulated experience, the ability to choose the substance, manipulation activities and the individual approach to the activity. The

creative performance has an abstract feature though it also refers to verbal, art, motor and music activities as the detailed layer of work.
The proposition scales were worked out in the framework of two indicators: a/ the relation of content of the text to the situation inspired by the creative process, b/ unconventional connection of content with specific structures. The results indicate a low level of children's creativeness. This is a sign of poor practice and calls for change.

The skill of classification

For the development a child needs sorted out knowledge, classified phenomena according to different criteria and the understanding of convention concepts. The scales refer to 3 indicators:

- criteria employed for ordering activities;
- the knowledge of convention concepts;
- the wise ordering activities. he research results indicate a low level of skills which are essential for development.

Conclusion

The presented proposition of estimating the educational effects in the category of development needs to be discussed. Different indicators and scopes can be applied in such a study. The problem needs to be elaborated, discussed, verified and changed in the educational process. The change needs to take place in the guidance through the children's creative activities and the abilities to work rationally.

USING EYE MOVEMENTS AND RETROSPECTIVE DATA TO STUDY THE USE OF CLEVER STRATEGIES IN CHILDREN'S MENTAL ARITHMETIC

L. Verschaffel, E. De Corte, I. Gielen, & E. Struyf, Center for Instructional Psychology and Technology, Laboratory of Experimental Psychology, University of Leuven, Leuven, Belgium

Background and research questions

In the present study eye-movement data and verbal reports were used to unravel whether or not 8-9-years-old children rearrange the order of the terms in horizontally presented two-step mental addition problems in order to shorten or simplify the solution process. Such rearrangements can be considered as "clever" ones, especially when somewhere in the problem an "easy" sum is hidden, such as the two complementary addends 3 and 7 in 3+8+7=. or two identical numbers as in 8+6+6=.

The goal of the study was twofold. From a *theoretical* point of view, we wanted to investigate whether good and weak mental calculators differ with respect to the use of the rearrangement strategy, and what task characteristics determine its application. But the study had also a *methodological* goal: By confronting pupils' eye-movements with their retrospective reports, we wanted to gather further insight into the strengths and the weaknesses of these two kinds of data for unraveling children's mathematical thinking.

Design

Based on a paper-and-pencil arithmetic speed test administered to a large group of third graders, 13 high-ability (HA) and 13 low-ability (LA) pupils were selected to participate in the experiment.

Each pupil was individually administered 48 target items that belonged to 6 types of number triples (a/b/c). These types differed with respect to:

- the equality of a and c: either a=c (e.g., 6/7/6) or a<>c (e.g., 6/9/4), and
- the size of the sum of a+c: either a+c>10 (e.g., 7/5/7), a+c=10 (e.g., 5/7/5), or a+c<10 (e.g., 4/9/4).

For each of the 6 types 4 concrete number triples were selected. The 24 selected number triples were presented in two different forms: one form in which a and c took the first and the third place (e.g., *6*+9+*4*=.), and one form in which they were placed in second and third order (9+*4*+*6*=.).

While the pupil was reading and solving the problems, his eye-movements were registered using DEBIC 80, a child-friendly system that uses the pupil center-corneal reflection method as its measurement principle. With this system, every 20 ms the x and y coordinates of the subject's point of regard, together with a time code, are stored. After the pupil nad solved the problem, he was asked to explain his solution strategy.
Pupils' raw eye-movement data for the 48 target items were reduced into consecutive fixations on the different problem elements. Starting from those fixation data, each solution process was scored into one of the following categories of solution strategies: a 12-, a 13-, or a 23-strategy, referring to a strategy that starts with adding the first and the second, the first and the third, and the second and the third given number, respectively. This scoring was based on the last fixation of at least 180 ms on one of the numbers; it was assumed that this last fixated number was added to the sum of the two other terms. In a separate analysis pupils' retrospective reports were also categorized as either indicating a 12-, a 23-, or a 13-strategy. Afterwards all 23- and 13-strategies were recoded as either an expected or an *unexpected* rearrangement strategy, depending on the place of the two "easy" numbers (e.g., for 4+7+6=. the 13-strategy is the expected rearrangement and the 23-strategy the unexpected one).
An analysis of variance was performed with ability level, equality of a and c, sum of a and c, and order of terms as independent variables. Dependent variables were: success rate, response time, use of the expected rearrangement strategy as evidenced by the eye-movement data, use of the expected rearrangement strategy as evidenced by the retrospective data, and correspondence between the strategy scores based on the eye-movement and the retrospective data.

Results

Theoretical results include the following. First, pupils made intensively and successfully use of the expected rearrangement strategy: two thirds of the retrospective protocols and half of the eye-movement protocols suggested the use of this strategy. However, we also found a significant amount of unexpected rearrangement strategies. Second, while HA pupils solved the problems better and faster than their LA peers, no significant difference in the use of expected rearrangement strategies was found. The three task variables, on the other hand, did have a considerable effect not only on problem difficulty and response time, but also on the use of expected rearrangement strategies. For example, a=c problems elicited more correct answers, faster response times, and more expected rearrangements than a<>c problems.
With respect to methodology, a rather good congruence was obtained between the strategy scores based on the subjects' eye-movement data and those derived from the pupils' retrospections: in about two thirds of the cases both scores indicated the same strategy (i.e., 12, 23 or 13). Interestingly, very high interindividual differences in congruence were observed. Although the detailed analysis is still ongoing, it seems that the incongruencies are mainly due to methodological and technical problems involved in the registration and interpretation of the eye-movement data, rather than to the unreliability of the pupils' retrospections.

Conclusions and Implications

The findings of this study show that already at a young age children make intensively and efficiently use of "clever" rearrangement strategies when doing mental addition. The data also confirm strong overall effects of particular task variables on the use of the rearrangement strategy. On the other hand, HA and LA pupils did not differ with respect to their use of rearrangement strategies. This latter finding suggests that the HA pupils' better performances and faster response times must be due to another factor, the most plausible candidate being that the HA pupils' had a more extended declarative knowledge base, and more specifically, their greater and more easy access to a large set of known facts. From a methodological point of view, we may conclude that the combined use of eye-movement and retrospection data can provide reliable and cross-validated information about children's overall strategies for solving mental arithmetic tasks. In our future work we will investigate whether the combination of these two techniques may also reveal more subtle aspects of these solution processes, which can not be obtained from verbal protocols alone.

LOGO AND PROBLEM SOLVING

W.van Dijk, ISOR, Department of Educational Research, Utrecht University, Utrecht, The Netherlands

Introduction

LOGO is of interest in education because this programming language offers various possibilities of learning through self-discovery. For example, children can send a turtle across the screen by means of various turn-and-move commands and instruct the turtle to make more or less complicated drawings.

LOGO is a procedural language. This means that a program can be split into smaller, independent procedures. In problem solving exactly the same thing happens: a problem is solved by being split into parts. In working with LOGO therefore pupils not only gain experience with (mathematical) concepts, but also learn to work with problem solving strategies.

Apart from being procedural, LOGO is also interactive in that each assignment can be executed immediately and the result is directly visible. This means that errors made in the procedures can be traced and corrected. This is known as 'debugging'.

The underlying idea behind LOGO is that it encourages metacognitive activities such as planning, debugging and dividing problems into parts. Research into the effects of learning with LOGO, however, has shown few positive results so far. Further examination shows, among other things, that LOGO instruction covers a restricted period in most investigations and that instruction is predominantly unsystematic. This criticism of research into the effects of LOGO led to the setting up of the 7005 SVO project entitled 'Effects of long-lasting LOGO instruction in primary schools' (Overtoom-Corsmit & Kunst, 1987). In this contribution we present a number of important findings from this research.

The following questions were addressed:

- What level do pupils reach in LOGO programming?
- Is the programming level determined by general IQ, cognitive style or negative fear of failure?
- Has working with LOGO a metacognitive effect or any effect on planning skills?

Design

Research variables and measuring instruments

In order to answer the research questions mentioned a quasi-experimental, pretest-posttest design was used with an experimental group (the 'logogroup') and

a control group. Figure 1 shows for each research question a survey of the research variables and the accompanying measuring instruments.

Figure 1: A survey of the research questions, variables and measurement instruments

research questions	variables	measuring instruments
What level have pupils reached in programming with LOGO?	programming skill	programming tests
Is the programming level determined by general IQ, cognitive style and negative fear of failure?	programming skill intelligence cognitive style negative fear of failure	programming tests Standard Progressive Matrices Embedded Figure test PMT-k
Has working with LOGO a metacognetive effect on the planning skill?	planning skill	Building Block test Tower of Hanoi test

Subjects

The pupils participating in the project (n = 118) were, at the beginning of the project, in year-groups 4 and 6 of four different primary schools . At two schools LOGO was used from groups 4 and 6 onwards for three consecutive years. The pupils from the other two schools formed the control group.

Procedure

Pupils in the experimental condition received an average of one hour a week of LOGO instruction (amounting to about 200 hours of LOGO instruction in total). The lessons took place at both schools in a separate classroom, which had been provided with ten home computers (Commodore-64). All tests were administered in written from to each group at the beginning of the research (Raven, Embedded Figures Test, PMT-k), as well as during the course of the project (programming tests, Building Block Test and Hanoi Test) and at the end of the research (again the Building Block Test and Hanoi Test).

Results

Programming levels

The results show that

- in the course of time pupils apply an increasing time of LOGO knowledge to their own programming;
- there is stagnation in the development of the programming level during the first and the last phases of the research;
- not all the knowledge targets are reached within the time limit;
  the youngest pupils (from group 4 to 6) lag behind the older pupils (from group 6 onwards) more than was expected.

The influence on the IQ, cognetive style and negative fear of failure

The answer to the question how far intelligence, cognetive style and negative fear for failure are on influence on the programming levels of pupils is different for each group examined. Multiple regression analyses show that the level of programming reached by the youngest group can only in parts be explained by the level of intelligence. In the older group (covering groups 6 to 8) part of the variance can be explained only by cognetive style. Taken in combination with intelligence and cognetive style, negative fear of failure has no predictive value.

planning skills

There are indicators that working with LOGO only had an effect on planning skills for the older group. We base these indications on the second administration of the Building Block test, in which the LOGO pupils in group 8 performed better than the control group pupils . At the first administration this difference between both conditions was still absent.

Discussion

Questions which promt themselves in the light of our investigation are the following:

- Why is the metacognitive effect of working with LOGO totally absent from group 4 onwards and not convincingly demonstrated from group 6 onwards?
- In spite of this lack of metacognitive effect, is the LOGO instruction at primary schools justified?
- How would LOGO instruction have to be structured in content and from and what from of research into its effects would be advisable?

References

Overtoom-Corsmit, R., & Kunst , H. (1987). Logo en probleemoplossings-vaardigheden. In: V. Haars (ed.), Leren met Logo. Nijmegen.

KNOWLEDGE BASED AUTHORING SUPPORT: A MODEL BASED APPROACH TO SEQUENCING

Anja van der Hulst, Department of Social Science Informatics, University of Amsterdam, Amsterdam, The Netherlands

Introduction

Instructional Design theory shows a tendency towards the development of finer grained theories. Where authors took advantage of rather general prescriptions, knowledge based authoring tools require but also enable the use of far more detailed theories.

This study describes the theory development for a tool that supports subject matter sequencing. A first implemented version of the theory incorporates a small number of prototypical models of sequences accompanied by a set of rules that prescribe which model applies when. These rules cope with for sequencing essential factors that have not been encountered in other prescriptive theories, such as temporal and physical aspects of the domain and in addition to that, a number of interaction effects.

Current theoretical approaches

Instructional design theory includes a variety of theories for sequencing. Comparing the prescriptions of these theories with the inventory of Posner and Strike (1976) who list aspects of domains and those of learning that have been found to be ground for sequential decisions made by teachers, currently available approaches ignore both characteristics of domains and aspects of the learning process. In particular Progressive Differentiation (Ausubel, 1963), the Spiral Curriculum approach (Bruner, 1966) and the Elaborative approach as proposed by Reigeluth and Stein (1983) are based on the idea of ``translating" the domain into some kind of subsumption hierarchy. With respect to the order in which elements should be presented, they prescribe to descend the hierarchy, which thus in all cases results in a sort of general to specific or deductive sequence. Other characteristics of the domain, such as possible temporal, physical or topological aspects, are ignored as a basis for decisions about sequence. Analyses of expert teacher behaviour however show (see e.g. Stevens, Collins, & Goldin, 1982) that many sequences are actually consistent with the way phenomena occur in the real world.

Besides, most theories prescribe only a single best approach for sequencing. However, Lodewijks (1981) and van Patten, Chao, and Reigeluth (1986) amongst others mention significant interaction effects of the factor "student intelligence" in

studies that measure the effects of inductive versus deductive sequences. In summary, available theories insufficiently identify and model sequencing variables and interaction effects.

Towards a finer grained theory

Models of sequences

In constructing a more detailed theory, a knowledge engineering approach, i.e. analysing expert teacher sequencing, was adopted. An attempt was made to identify and model prototypical forms of sequences.

To start with, examples were sought of instructional material that covered the classes of domains characteristics as listed by Posner and Strike. For the analysis of this material, an instrument for the analysis of written instruction previously developed (van der Hulst & de Jong, 1991) was used. By means of this instrument, a collection of models of tactics, an interpretation of instructional material could be made in terms of the domain elements taught and the tactics and by that strategies used.

The results will only be illustrated by means of an analysis of the sequence of topics of four lessons covering the subject of 'validity' in a course about statistics and methodology. The lessons showed an (approximately) identical hierarchical domain structure. The most inclusive concept was that of 'validity' with as major types, 'content validity', 'predictive validity' and finally 'construct validity'. In the majority of books of each of the major type again a number of subtypes were described.

An remarkable finding was that way the different authors sequenced the material was extremely similar. The analyses consistently showed a top-down (general-specific) - depth first sequence. Similar sequences were found in lessons covering the subject of kinematics, whereas breadth first patterns were observed in sequences covering causal branched chain like structures.

The former analyses led to a first basic set of models of prototypical sequences. The main idea that underlies this basic set of models is that a very limited number of basic structure forms could be defined, i.e. linear structures and tree, or branched chain, structures (other forms, such as cyclic of web structures being either a kind of- or composed of the basic forms). For each of the basic forms again only a very limited number of 'natural' sequences could be specified. Features of these models of sequences refer to the structure form they apply to, a global direction and a pattern. As such, the theory incorporates for instance a sequence model for a `tree' structure following a top-down direction using a depth-first pattern.

The prescriptive part

Its support however goes further. The theory prescribes which sequencing model for the certain structure is assumed to be most appropriate, considering the nature of the relations in the structure, and prior instructional decisions. To illustrate this, in van Patten, Chao, and Reigeluth (1986) a number of studies have been reviewed that are rather consistent about the effects of both example-generality and generality-example sequences. Each of the sequences results in different outcomes, a generality-example sequence appeared to lead to better use -a generality (or awareness of meanings of concepts) outcomes, whereas an

example generality or "eg-rule" sequence results in better find-a-generality outcomes, i.e. better categorisation. These findings might give rise to defining a prescriptive rule that, given the selected instructional goal (being either use-, or find a generality), both prescribes the strategic approach (e.g. exposition or inquiry) and the sequence (top-down, bottom-up) to be selected.

The authoring tool

A prototype authoring tool that implements a first and incomplete version of the theory is able to decompose the usually hybrid subject matter structures into primitive ones for which models of sequences are available. Having analysed the relation type within the structure, it suggests one or more 'natural' sequences. Finally, since it was assumed essential to provide authors with a justification of the sequence selection, the system intergrates a theory inspection mode. The inspector for instance shows the relation types for which the system includes prescriptions and provides on request the sources (e.g. observations on expert teachers, comparative studies) of the prescriptions.

Acknowledgements

The study described here is part of a project called MIOS and is financed by the University of Amsterdam in its "Profilerings convenant". Hereby I'd like to acknowledge Paul Kamsteeg, Ton de Jong, Jacobijn Sandberg, Yvonne Barnard, Lynda Hardman, Robert de Hoog and Bob Wielinga for discussions and comments on topics reported in this paper.

References

Ausubel, D. (1963). The psychology of meaningful verbal learning: An introduction to school learning. New York: Grune + Stratton.

Bruner, J. (1966). Toward a theory of instruction. New York: W.W. Norton.

Hulst van der, A., & Jong de, T. (1991). Een reconstructie van instructie: ontwikkeling van prototypische modellen van onderwijsacties. [A reconstruction of instruction: development of prototypical models of instructional actions]. In J. Beishuizen (Ed.), ORD' 91: Technologie (pp. 17-26), Amsterdam: VOR, SCO.

Lodewijks, J. (1981). Leerstofsequenties van conceptueel netwerk naar cognitieve structuur. [Subject matter sequences from conceptual network to cognitive structure]. PhD thesis, University of Utrecht.

Posner, G., & Strike, K. (1976). A categorisation scheme for principles of sequencing content. Review of educational research, 46, 665-690.

Reigeluth, C., & Stein, F. (1983). The elaboration theory of instruction. In C. Reigeluth (Ed.), Instructional design theories and models, (pp. 335-379). Hillsdale, New Yersey: Lawrence Erlbaum Associates.

Patten, van J., Chao, C., & Reigeluth, C. (1986). A review of strategies for sequencing and synthesizing instruction. Review of educational research, 56(4), 437-471.

GENDER RELATED DIFFERENCES IN MOTIVATION AND ARITHMETIC ABILITY

Gerard Seegers, & Monique Boekaerts, Centre for the Study of Education and Instruction, Leiden University, Leiden, The Netherlands

In several studies it was found that boys have a superior math ability (cf. Meyer & Fennema, 1988). These gender-related differences seem to be present from an early age onwards. Recently, this pattern was confirmed by the results of the large-scale evaluation study conducted by the National Institute for Test Development (CITO). Wijnstra (1988). demonstrated that at the primary school level, differences between boys and girls are least apparent in relation to algorithmic tasks, and are most manifest in relation to mathematical problem solving. There is no easy explanation for this pattern of results, but there is growing consensus that a dominantly cognitive point of view is inadequate. Today, researchers believe that gender-related differences in math are the outcome of complex interactions between cognitive abilities and attitudinal and motivational variables. For example, a study by Nicholls, Cobb, Wood, Yackel and Patashnick (1990) displayed that students who are ego-oriented ascribe success in mathematics primarily to superior ability, while students who are task-oriented are more inclined to emphasize effort when achieving success.

The constructs of ego and task orientation were proposed by Nicholls (1983) to refer to the students' goal orientation. He documented the view that pupils who are task-oriented are more interested than ego-oriented pupils in the learning tasks they have to do, and are more willing to invest effort in completing them. They tend to evaluate their results positively, as long as they perceive improved performance. On the other hand, ego-oriented pupils are less interested in learning tasks, and have a more competitive attitude. Their willingness to invest effort in learning tasks is dependent upon the probability of achieving success, i.e. doing better than the reference group. They will loose interest in the task if they perceive a task as over-taxing their capacities, thus fearing that failure will be seen as lack of capacity (see also Dweck, 1975). As far as sex differences are concerned, there is some evidence that male students are more inclined to attribute success to ability and display more interpersonal competition (Spence & Helmreich, 1983).

Goal-orientation and attributional style

In the present study we wanted to investigate whether the findings reported by Nicholls and his co-workers could be replicated in a Dutch sample of 6th graders (pupils around the age of twelve). More specifically, we hypothesized

- that boys would be more ego-oriented than girls, and
- that ego-oriented pupils would attribute success in math primarily to superior ability (stable, uncontrollable attribution), and
- that task-oriented pupils would emphasize effort attributions (variable, controlable attribution) after perceived success.

In addition, we wanted to investigate whether situation-specific cognitions about math tasks would help explain the variance in math achievement. The measurement instruments used in this study were an adapted version of Nicholls' ego- and task-orientaion scale, and an adapted version of Weiner attribution scale (cf. Weiner, 1985). Arithmetic ability was assessed by means of a number of items from the large scale Cito-math evaluation study. These items cover a broad range of specific subjects within the arithmetic curriculum, and measure specific math abilities. Items range from simple algorithmic ciphering to complex mathematical problem solving in a realistic context. Boekaerts (1985, 1987) theorized that pupils perceive differences in learning tasks and that differential perception of learning tasks is reflected in their situation-specific judgments of the task as well as in the learning intention. She constructed the on-line motivation questionnaire (OMQ) with which pupils' cognitions and feelings about a learning task can be measured. The OMQ was administered twice to a group of 6th grade pupils (N= 221). The students made the math tasks in two separate sessions of one hour. These tasks were matched in order to present pupils with two comparable set of learning tasks. At on-set and off-set of these two sets of math tasks pupils filled in the OMQ.

The results of this study confirm the findings reported by Wijnstra: Boys displayed a higher score on the arithmetic test than girls. And, differences between boys and girls grew larger as the complexity of the items increased. In the second place, our results support the findings reported by Nicholls and his co-workers: Girls more often ascribe failure to lack of ability than boys, while boys score higher on the the most central construct of ego-orientation, as conceptualized by Nicholls, viz. 'demonstrating ability' (but not on the two other subscales of ego-orientation viz. 'hiding mistakes from others' and 'fear of making mistakes'). We also found that for both boys and girls, task-orientation was positively associated with the tendency to ascribe success to effort. In boys, task orientation was negatively related to the tendency to ascribe failure to lack of ability, but in girls no significant relation was found. Differences between boys and girls were also observed in the relation between ego-orientation and their 'attribution of failure to lack of ability'. Boys who scored high on the central ego-orientation subscale 'demonstrating ability' scored low on 'attribution to lack of ability'. This relation was non-significant for girls. By contrast, girls who scored high on the two other subscales of ego-orientation: viz. 'hiding mistakes' and 'fear of making mistakes' displayed frequent attribution of failure to lack of ability. For boys this relation was non-significant. Finally, no sex differences were displayed in the relation between ego-orientation (3 subscales) and 'attribution of success to effort': all relations were non-significant. This pattern of results seems to indicate that ego and task orientation have a different relation to attributions for boys and girls.

Goal-orientation and arithmetic performance

We also wanted to study the *direct effects* of goal-orientation, and attribution style on

- situation-specific cognitions about the math tasks, and
- arithmetic performance, and the *indirect effects* via situation-specific cognitions and learning intentions on arithmetic performance. It was predicted that goal-orientation and attributional style would influence arithmetic performance via situation-specific cognitions. More specifically, we hypothesised that ego and task orientation would differentially affect the appraisals measured by the OMQ (subjective competence judgments, and intrinsic pleasure judgments) and learning intention (also measured with the OMQ), and that this would be reflected in arithmetic performance. A parallel prediction was formulated for attributional style. Gender differences were also expected. Tests for nonverbal intelligence and self-efficacy were also administrated and included as variables in the model. This model was tested with the LISREL programme. The results show that ego-orientation, in the form of 'hiding mistakes from others', and 'tendency to attribute failure to lack of ability' contributed negatively to the students' subjective competence judgment. A positive relation was found between task-orientation, intrinsic pleasure judgment, and learning intention. Interestingly, a strong effect of self-efficacy was demonstrated on both subjective competence judgment and arithmetic performance. The results revealed differences between boys and girls. These differences concern the position of *subjective competence* and *intrinsic pleasure judgement*. As far as subjective competence is concerned, we noticed a strong influence of goal-orientation and attributional style in girls, while the influence of self-efficacy was found to be more dominant in boys. Inspecting intrinsic pleasure judgement, in girls it was found to be strongly influenced by subjective competence, while in boys the influence of task- orientation was found to have a more dominant influence. The implications of these results will be discussed.

References

Boekaerts, M. (1985). Some new developments in the study of motivational processes in a classroom context. In G. D'Ydewalle (Ed.), Cognition information processing and motivation (Vol. 3, pp. 569-586). Amsterdam: North Holland.

Boekaerts, M. (1987). Die Effekte von State- und Trait-Motivationaler Orientierung auf das Lernergebnis. Zeitschrift für Pädagogische Psychologie, 1(1), 29-43.

Dweck, C.S. (1975). The role of expectations and attributions in the allevation of learned helplessness. Journal of Personality and Social Psychology, 31, 674-685.

Fennema, E. (1985). Attribution theory and achievement in mathematics. In S.R. Yussen (Ed.), The Growth of Reflection in Children (pp. 245-265) Orlando: Academic Press.

Meyer, M.R., & Fennema, E. (1988). Girls, boys, and mathematics. In: T.R.Post (ed.) Teaching mathematics in grades K-8: Research-based methods. (p. 406-425). Newton,MA: Allyn & Bacon.

Nicholls, J.G. (1983). Conceptions of ability and achievement motivation. In S.G. Paris, G.M. Olson, & H.W. Stevenson (Eds.), Learning and Motivation in the classroom. Hillsdale, NJ: Erlbaum.

Nicholls, J.G., Cobb, P., Wood, T., Yackel, E., & Patashnick, M. (1990). Assessing students' theories of success in mathematics: Individual and classroom differences. Journal for Research in Mathematics Education, 2(12), 109-122.

Spence, J.T., & Helmreich, R.L. (1983). Achievement-related motives and behaviors. In: J.T.Spence (ed.). Achievement and Achievement Motives (p.7-74). San Francisco: Freeman

Weiner, B. (1985). An attributional theory of achievement motivation. Psychological Review, 92, 548-573.

Wijnstra, J.M. (Ed.) (1988). Balans van het rekenonderwijs in de basisschool: Uitkomsten van de eerste rekenpeiling medio en einde basisonderwijs. Arnhem: Cito (PPON reeks no. 1).

DUTCH REVISION OF HARTER'S SELF-PERCEPTION PROFILE FOR ADOLESCENTS

Nitha M.E. Neuwahl, & Coosje Groen-de Jong, Leiden University, Leiden, The Netherlands

Introduction

Positive self-concepts and positive self-esteem are important for students' psycho-social functioning as well as for their learning achievements (Heckhausen, 1987; Pope et al., 1989). If one wants to gather knowledge of students' self-concepts and self-esteem, it is important to have reliable measures for these concepts. One example of an adequate American instrument is Susan Harter's Self-Perception Profile for Adolescents (1986). Harter's theoretical framework goes back to William James (1890). For adolescents Harter distinguishes eight domain specific self-concepts: Scholastic competence, social acceptance, athletic competence, physical appearance, behavior/morality, job competence, romantic appeal, and close friendship. General self-esteem has a separate place apart from these domain specific self-concepts. As some other researchers (e.g. Meyer, 1984) Harter states that general self-esteem is not a simple adding-up of domain specific self-concepts. The extent to which a domain specific self-concept contributes to the general self-esteem depends on the extent to which that self-concept is important for the person.

Dutch Revision of Harter's questionnaire

In a former study (Neuwahl & Groen-de Jong, 1991) we translated Harter's questionnaire for adolescents into Dutch. Analyses of our data showed differences compared to Harter's analyses: The structure of the questionnaire was not the same; not all eight self-concept domains could be found in our study. Another difference was that we could construct reliable scales only out of some of the factors. Because of these non-optimal qualities of the translated questionnaire, and because of the fact that Harter reformulated a number of her items (1988), we made a revised version of the questionnaire. In this revision we reformulated the non-optimal items and we added items on 'relationship with parents', a scale which Harter thinks important for college students and which we think also important for adolescents. We not only made this revision of Harter's Self-Perception Profile for Adolescents, but we also made a revision of Harter's questionnaire that measures the importance of the domain specific self-concepts for a person's general self-esteem. In this last questionnaire each domain has only two items. Because in a former study (Caron-Derkits, 1987) the internal consistencies of the scales were

376

too low, we added items in order to try to improve the quality of this questionnaire. Each domain in both revised scales consists of five items.

Subjects

The revised Dutch versions of Harter's questionnaires were filled in by 749 students in different levels of secondary education (358 boys and 391 girls), Grades 7 through 10/11 (226 Grade 7, 192 Grade 8, 166 Grade 9, 140 Grade 10, and 25 Grade 11). They came from two schools in the centre of the Netherlands.

Results

Factoranalyses and internal consistencies of the scales

We studied the structures of the questionnaires. They were very much like those Harter described. In the Self-Perception Profile for Adolescents six factors were identical: scholastic competence, athletic competence, physical appearance, behavior/morality, job competence, and relationship with parents; all items of two of Harter's factors, social acceptance and close friendship, loaded on one factor in our study; the items of one factor, romantic appeal, loaded on two factors in our study. The internal consistencies of the scales, constructed by means of unweighted adding-up of the items that loaded .30 or higher on the factors, were greater than .70 except for 'scholastic competence and for one of the factors 'romantic appeal'. Factoranalysis of the questionnaire on the importance of the self-concepts resulted in the same factors as was a-priori described: all factors consisted of the five items which were supposed to load on the factor, with one exception: two items of 'romantic appeal' loaded on 'close frienship'. The internal consistencies of the scales, constructed by means of unweighted adding-up of the items that loaded .30 of higher on the factors, are > .70 for all scales except for the one for 3-items scale 'romantic appeal'.

Scores

Scores were calculated for scales that resulted from the factoranalyses and had internal consistencies of .70 or higher. Because we wanted to compare the results from both questionnaires, we decided to present here the data for the scales athletic competence, physical appearance, behavior/morality, job competence, relationship with parents, and social acceptance including close frienship. Scores were also calculated for general self-esteem, which, according to Harter, has a seperate place apart from the domain specific self-concepts and therefore its items were left out in the factoranalyses. Score were calculated for the total group, for both sexes, and for different age groups and grades. (Because there were only 25 students in Grade 11, those were put together with the students in Grade 10 in the group Grade 10/11.) The results for the different ages are very much like those for grades; we present those for grades.

Two-factorial 2 x 4 (sex x grade) ANOVAS were performed separately on the six scales. They revealed main effects for sex on five scales: Social skills, $F(1,644)=13.30$, $p=.000$; physical appearance, $F(1,644)=30.22$, $p=.000$; athletic competence, $F(1,644)=106,49$, $p=.000$; job competence, $F(1,644)=9.19$, $p=.003$; and behavior/morality, $F(1,644)=13.17$, $p=.000$. Boys have higher scores than girls on physical appearance, on job competence, and on athletics; girls have higher scores than boys on social skills: social competence and close friendship, and on

377

behavior/morality). In relationship with parents no sex differences were found. For job competence there was also a grades effect, $F(3,644)=11.09$, $p=.000$, girls have higher scores than boys, except in Grades 10/11). For the other scales no grades effects were found. An interaction effect, sex x job competence, $F(3,644)= 4.75$, $p=.003$ was found. This interaction effect is probably due to the fact that for boys scores are decreasing after Grade 7 and increasing after Grade 8, while scores for girls are also decreasing after Grade 7, but are further decreasing in Grades 8, 9 and 10/11.
Also for the importance of the domain specific self-concepts two factorial ANOVAS were performed. They revealed main effects for sex on all scales: Social skills, $F(1,657)=42.79$, $p=.000$; physical appearance, $F(1,657)=15.41$, $p=.000$; athletic competence, $F(1,657)=79.11$, $p=.000$; job competence, $F(1,657)=16.74$, $p=.000$; relationship with parents, $F(1,657)=23.67$, $p=.000$; and behavior/morality, $F(1,657)=9.60$, $p=.002$. Girls think the importance of the self-concepts for the general self-esteem greater than boys do, an exception is that boys think athletics more important than girls do. Main effects for grades were found for the importance of athletic competence, $F(3,657)=6.22$, $p=.000$, and for the importance of behavior/morality, $F(3,657)=5.93$, $p=.001$. Scores on both scales are decreasing with increasing grades. An interaction effect, sex x grade, was found for the importance of physical appearance, $F(3,657)=5.01$, $p=.002$. For boys there is a drop in scores after Grade 7, scores in Grades 8 through 10/11 had about the same height, while for girls the scores are increasing with inceasing grade.

Correlations
Correlation between the domain-specific self-concepts and general self-esteem were calculated. The highest correlation (.66) was with physical appearance. Medium high correlations were found with relationship with parents (.39), with social skills (.36), with behavior/morality (.25), and with athletic competence (.23). The correlation with job competence was still significant at 1% level, but was only .12.

Conclusion and discussion
It can be concluded that the revised instruments have good psychometric qualities. The two scales that are non-optimal (scholastic competence because the internal consistency is <.70, and romantic appeal because the items load on two separate factors in our study) were improved in a new study. In that study, from which not yet all results are available, we were able to construct separate scales for scholastic competence and for romantic appeal with internal consistencies of .78. We now have available adequate revisions in Dutch of Harter's questionnaires for adolescents.
The groups Harter studied are not fully comparable to ours, but as far as they are comparable, the results are very similar. We found, as Harter did, higher scores for boys than for girls in athletic competence, in physical appearence and in job competence, and higher scores for girls than for boys in social skills (social acceptance and close friendship). Harter gives no results on behavior/morality. Harter presents, as we have found in many of our studies, lower self-esteem for girls than for boys, especially in the higher grades. The correlations between the domain specific self-concepts and general self-esteem are also similar. The

highest correlation Harter found (.71) was, just as we found, with physical appearance. For the other scales she found correlations from .33 to .51; job competence has a correlation of .38, higher than in our study. Obviously American adolescents are in many aspects of their self-concepts, but not in all, comparable to Dutch adolescents.

References

Caron-Derkits, I. (1987). Over de waarneming van de eigen bekwaamheid bij kinderen van het lager onderwijs. Een exploratief onderzoek naar de bruikbaarheid van het 'Self-Perception Profile for Children' voor een Nederlandse populatie van kinderen in de basisschoolleeftijd. Leiden: Rijksuniversiteit, vakgroep Ontwikkelingspsychologie.

Harter, S. (1986). Processes Underlying the Construction, Maintenance and Enhancement of the Self-Concept in Children. In: J Suls & A.G. Greenwald (Eds.). Psychological Perspectives on the Self. 3. New Jersey: LEA.

Harter, S. (1988). Manual for the Self-Perception Profile for Adolescents. Denver: University.

Heckhausen, H. (1987). Causal attribution patterns for achievement outcomes: Individual differences, possible types and their origins. In: F.E. Weinert & R.H. Kluwe (Eds.). Metacognition, Motivation and Understanding. Hillsdale: Erlbaum.

James, W. (1890). The consciousness of the self. In: James, W. The Principles of Psychology. 1. New York: Holt & Co.

Meyer, W.-U. (1984). Das Konzept von der eigenen Begabung. Bern, Stuttgart, Toronto: Huber.

Neuwahl, N.M.E., & Groen-de Jong, A.C. (1991). Zelfbeeld en zelfwaardering van adolescenten. Onderzoek naar de bruikbaarheid van het meetinstrument van Susan Harter. In: B.H.A.M. van Hout-Wolters & L.F.W. de Klerk (Eds.). Onderwijsleerprocessen: cognitie en motivatie. (33-42). Amsterdam: SCO.

Pope, Alice, W., McHale, Susan M., & Craighead, Edward W. (1989). Werken aan het gevoel van eigenwaarde. Assen: Dekker & van de Vegt.

INTEGRATING THE AFFECTIVE DOMAIN INTO THE INSTRUCTIONAL DESIGN PROCESS

Robert G. Main, California State University Chico, USA

Introduction

This research investigates how the affective domain can be addressed systematically in the instructional design process. The goal is to produce an integrated model of instructional design that includes the affective domain as an essential component.

For purposes of this paper, the affective domain is characterized in terms of motivation as it affects the direction and intensity of behavior. This is not meant to be restrictive of the factors involved in establishing motivation, but rather as a convenience in bundling the many components and how they are considered in relation to including the affective domain in instructional design. This definition reflects Gagne's (1988) concept of attitudes and Keller's (1979) concept of motivation in developing instructional plans and activities designed to influence the learner's effort to achieve a desired performance. It does not preclude the consideration of aptitude and trait interactions as determinants of learner performance, but they would be considered only in terms of their contribution to motivating the learner's behavior.

Neglect of the Affective Domain

Motivation is the neglected "heart" of our understanding of how to design instruction, according to Keller (1979). "Historically, instructional science has benefitted from the work of behavioral psychology and cognitive-learning psychology, but this has given us only partial knowledge of how people learn, and almost no knowledge of *why* they learn" (p. 390). According to Beane (1985/86) "... the form or substance of affective education represents perhaps the most problematic of all school issues" (p. 27).

Technology itself may be a contributing factor in the neglect of affective learning objectives. Computer based training (CBT) has been the leading edge for instructional technology for almost three decades. It is an expensive technology and education and training administrators are sold on innovation by cost/benefit analysis. Consequently, the pressure for each has been to determine how much and how fast knowledge could be acquired and skills attained using CBT. Besides there is something Orwellian about having a computer teaching attitudes and values—especially in the public schools. Keller (1983) states, "... we often read that the goal of instructional technology is to design effective and efficient

instruction. Unfortunately, these criteria make it easy to *exclude a specific concern for motivation*, or the appeal of instruction: (p. 388).

The Integrated Instructional Design Model

Affective domain instruction can be divided into two areas. One of these deals with instruction where the subject matter itself is principally concerned with changing student values, beliefs and attitudes. Courses in race relations, ethics and drug abuse prevention fall into this category as do classes in a nation's history and traditions that are primarily concerned with generating a sense of loyalty and pride. The second area of affective domain instruction addresses how the learner feels about the subject being learned. The goal is simply to motivate the learner to *want* to master the knowledge and skills being taught. We need to spend as much effort in motivating the student to learn in the design and delivery of instruction as we do with the cognitive and psychomotor needs. Perhaps we should spend more time and attention since it has such a powerful impact on achievement.

It is evident from the literature that the affective domain is an important area in education and training—both in terms of achieving affective behaviors and the facilitation of cognitive and psychomotor objectives. The development of clearly defined instructional activities and strategies for the affective domain has lagged those of the psychomotor and, particularly, the cognitive domains. Current instructional system design models have been developed principally for use in developing instruction for cognitive objectives.

To correct this problem and insure that affective domain objectives are addressed in every instructional lesson, the affective component of instruction must be embedded within the instructional system design model. This paper presents an instructional design model that integrates the work of Keller (1983) in his model for motivating the learner with the five phase ISD model in general use by industry and military trainers. The model proposed will make sure the affective domain is considered in a systematic and routine fashion from curriculum planning and design through lesson development, delivery and evaluation of learning outcomes.

The A R C S Model

Keller has developed a general model integrating the various sources of motivation for learning. He refers to it as the ARCS model; an acronym for the four sets of conditions that must be met in order to have a motivated learner: A for attention; R for relevance; C for confidence; and S for satisfaction.

Attention involves grabbing the learner's interest at the beginning of instruction and maintaining that interest throughout the lesson and the course. *Attention* sustaining events arose the learner's curiosity. *Relevance* is the personal significance and value to the learner of mastering the learning objectives. The most straightforward tactic, according to Keller, is to reform the learner importance of the learning outcome to some desired state or goal, i.e., completing a technical course will provide eligibility for a promotion. The point is that the goal is desirable from the learner's perspective—not the lesson developer. *Confidence* relates to the learner's expectancy of success. Keller maintains that personal expectancy for success is influenced by past experiences (success or failure at the given task) and locus of control and personal causation (personal control and competence).

Difficulty of tasks are also a factor. Success at simple tasks may not generate confidence. *Satisfaction* is derived from achieving performance goals. The gratification of goal achievement is confounded to some degree by whether the evaluation of learning outcomes are externally based or made by the learner internally. Keller speculates that because heavy doses of performance evaluation characterize instructional design, it is not difficult to see that as part of the reason for the erosion of the intrinsic interest of children in the school process.
Keller makes a point to distinguish between effort and performance as factors in motivation. He sees effort as the primary dependent variable of motivation while performance is influenced by ability (individual characteristics) and opportunity (learning design and management) and only indirectly related to motivation. A further distinction is made between performance and consequences. Consequences include affective responses, social rewards and material objects. Consequences combine with cognitive evaluation to influence changes in personal values or motives. Affective behavior is considered to be a function of both person and environmental factors.

The Integrated Affective Domain/ISD Model
A conception of how Keller's ARCS model can integrated with a modified version of the military ISD model (Tennyson, 1989) to create a matrix of the design process is shown in Figure 1.

Instructional design phases

AFFECTIVE DOMAIN	Analysis	Design	Development	Implementation	Evaluation
Attention					
Relevance					
Confidence					
Satisfaction					
	Validation Feedback				

Figure 1
ISD Model Integrating Affective Domain
Across the top of the model are the five phases of the military ISD model: *Analysis, Design, Development, Implementation, and Evaluation.* Down the left side of the figure are the four categories defined by Keller as components of motivation: *Attention, Relevance, Confidence, and Satisfaction.* Along the bottom of the model is a rectangular cell labelled *Validation/Feedback* to depict the formative evaluation which occurs throughout the instructional design process and serves to validate that the tasks in each phase have been completed and reviewed. The arrows indicate the two-way flow of information between the phases that provide

382

feedback for improvement and maintenance of the system. It also shows the process is on-going and not necessarily linear.

This is an attempt to provide a first tentative step in addressing affective components within the instructional design process. To become a useful tool for instructional designers and lesson developers, the model must be fleshed out with task lists and taxonomies of strategies for each cell.

The theories of instruction and models of instructional design focus by and large on the cognitive and psychomotor domains. The affective domain is recognized by most in the literature, but in practice is largely ignored as an area of scientific research in the instructional technology field.

A great deal has been written about the changing role of teachers in computer based instruction. They are to become more managers of instruction than the presenters of instruction and the instructor/student ratio increases in the computer based training environment (Kearsley, 1983). As that occurs, the role shifts more and more to manager/technician as the principal duty becomes keeping the technology on line and managing the instructional environment. In this scenario, the affective domain receives less and less attention unless the art of teaching is incorporated in the CBT lessons. The way to insure the affective domain is given consideration and treated systematically in all instructional environments is to have it embedded within the instructional design model.

The lack of explicit depiction of the affective domain in contemporary instructional system design models does not necessarily mean that instructional developers do not include affective considerations in their lessons. It is simply no possible to design either cognitive or psychomotor instruction without including some affective component. The very act of establishing an instructional goal implies some value to the person, organization or society in its achievement. The selection of content to be included or excluded in the lesson requires the exercise of judgment as to the importance or worth of the knowledge and skills to be taught.

It is precisely because the affective is so entwined with cognitive and psychomotor learning achievements that it needs careful and conscious attention during the design and development of instruction. That is why this model should be helpful. It provides that the *systematic* consideration of affective objectives be integrated into every aspect of the instructional design process.

References

Beane, J. A. (1985/86). The continuing controversey over affective education. Educational Leadership, 43 (3) p. 26-31.

Gagne, R. M., & Driscoll, M. P. (1988). The essentials of learning for instruction. 2d ed. Englewood Cliffs. New Jersey: Prentice-Hall.

Kearsley, G. (1983). Computer based training: A guide to selection and implementation. Reading Massachusetts: Addison-Wesley.

Keller, J. M. (1979). Motivation and instructional design: A theoretical perspective. Journal of Instructional Development, 2 (4), p. 26-34.

Keller, J. M. (1983) Motivational design of instruction. In C. M. Reigeluth (ed.) Instructional-design theories and models. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Tennyson, R. D. (1989). Cognitive science update of instructional systems design models. Technical Report. Lexington, Massachusetts: Mei Associates.

THE SELF-REGULATED LEARNER IN SOLVING MATHEMATICAL PROBLEMS

Maria Cardelle-Elawar, Arizona State University West, USA

Introduction

The present study describes and illustrates the practical applications of metacognitive instruction applied to Mayer's model to solve mathematical problems. This approach has been tested empirically in three studies and the results indicated that low-performer students can be self-regulated learners when solving mathematical problems.

Theoretical Framework

According to the Standards for School Mathematics, prepared in 1989 by the National Council for Teachers of Mathematics (NCTM), problem solving and critical thinking should be the central focus of mathematics instruction. Classrooms with problem-solving orientations should be permeated by thought-provoking questions. Achieving this objective presents special challenges and opportunities for teachers of students who are educationally disadvantaged.

Low-performing students commonly exhibit counter-productive habits in approaching mathematics problems:

- They read rapidly at the expense of full comprehension of the question;
- they lack the skills to reorganize information;
- they are not sufficiently careful in recognizing that there might be more than one right way to solve the problem;
- they are uncertain about how to calculate and verify the solution (Anderson, 1990; Cardelle-Elawar, 1990; Romberg & Carpenter, 1986; Silver, 1981).

Metacognition

Metacognition is a systematic strategy for solving problems that includes reflecting on and evaluating the productivity of one's thinking. The key aspects of metacognition are self-monitoring, self-regulating, and the developing of an awareness of what one needs to know to solve problems (e.g.,Brown, 1987 Haller, Child, & Walberg, 1988; Resnick & Klopfer, 1989).

Mayer's Model for Solving Mathematics Problems

Students were instructed to solve problems following Mayer (1987, 1985) model (see Figure 1). Mayer suggested that four types of *processes* or knowledge are required to solve mathematics problems: translation, integration, planning and

monitoring, and solution execution. *Translation* requires good declarative and procedural knowledge. It often is difficult to translate a problem from its linguistic representation to a mental representation. *Integration* requires the student to combine problem statements into a coherent representation. In order to integrate the information in problems, students need the schematic knowledge to recognize problem types. (Schoenfeld, 1985, 1989; Silver (1987) found that good problem solvers possess more useful schemata than poor problem solvers. *Planning and monitoring* require knowledge of strategies that focus on how to solve the problem. The plan involves breaking the problem into subproblems and establishing a sequence for the solution (e.g., What operation should be employed first, second, etc.?). *Solution execution* requires the student to use procedural knowledge to apply the rules of arithmetic accurately and efficiently while carrying out the calculations in the solution plan (e.g., 65.2 - 12.1 = 53.1). Devising and monitoring a solution plan are crucial to mathematical problem solving.

Results

Analysis of variance performed on 500 students from elementary classrooms (240 in the experimental group and 260 in the control group) indicated that experimental group performed significantly higher than the control group. The findings indicate that low-mathematics ability students progressed as problem solvers in (a) *understanding* how to approach a problem, (b) *identifying* the appropriate schema for organizing the information, (c) *recognizing* there may be more than one right way to solve the problem, and (d) *verifying* their solutions. Understanding how to approach a problem was improved by increasing their linguistic 2comprehension concerning the meaning of key words and sentences. This improved comprehension heightened their concentration on the problem and reduced their impulse to say just anything for the sake of making a response. The questions derived from Mayer's model directed attention to and stimulate discussion of multiple ways to solve a problem as evidenced by the students' new ability to describe their solution process and compare their process to those used by other students. Students learned to verify their solutions by reflecting on their own thinking and examining the steps they took in working out the solution while checking the accuracy of calculations.

Recommendations

These study findings suggest several recommendations for teachers trying to help low performers in mathematics. First, the focus should not be on the students' labels (e.g., low performers, minority, etc.) but on each student's individual behavior. Special consideration should be given to each student's individual uniqueness, strengths and weaknesses. Second, low-performing students need a supportive atmosphere in which errors and mistakes lead to positive feedback and direction. These results provide some evidence to change the misconception that only bright students are able to do mathematics. Third, low-performing students need much structure in the classroom. Having the task broken down into small doses of learning and mastery seemed to help. A major component of the necessary classroom structure is that the problem presentation be well organized and structured. The academic task presented the student is a major influence on what the student learns. The key to a student's learning lies in his/her own

thoughts and actions. To learn, the student must practice thinking and not just apply procedures in a rote fashion. Teachers have traditionally used rote methods emphasizing skills in part because they are faster and seem more efficient, even though most students do not understand much of what they are doing (Wood, Cobb, & Yackel,1991). Fourth, a great deal of interaction between teacher and student is required. Students learn best through involvement and through the teacher's mediation and use of constructive feedback. Teachers must develop the sensitivity to know when to intervene to make appropriate suggestions and how to allow students time to think on their own. Such techniques can make mathematics a source of enjoyment rather than frustration.

Finally, this study responds to the current reform in mathematics as advocated by the National Research Council (1989), NCTM, (1991) and Wood, Cobb & Wackel (1991), National Assessment of Educational Progress (1988a) (NAEP)Brown, Carpenter, Kouba, Linquist, Silver, and Swafford,(1988b) who call for a radical change in tradition. These advocates emphasize the conceptual nature of mathematics and the active role of students. In this study students profited from reflection. Contrary to the conventional wisdom that low performers lack metacognitive skills, these students demonstrated metacognitive potential when stimulated by explicit individualized instruction.

References

Mayer, E. R. (1987). Educational Psychology. Boston: Little, Brown and Company.

Mayer, E.R. (1985). Mathematical ability (pp.127-150). In R.J. Sternberg (Ed.), Human abilities: An information-processing approach. New York: Freeman.

National Council of Teachers of Mathematics. (1991). Professional standards for teaching mathematics. Reston, VA: National Council of Teachers of Mathematics.

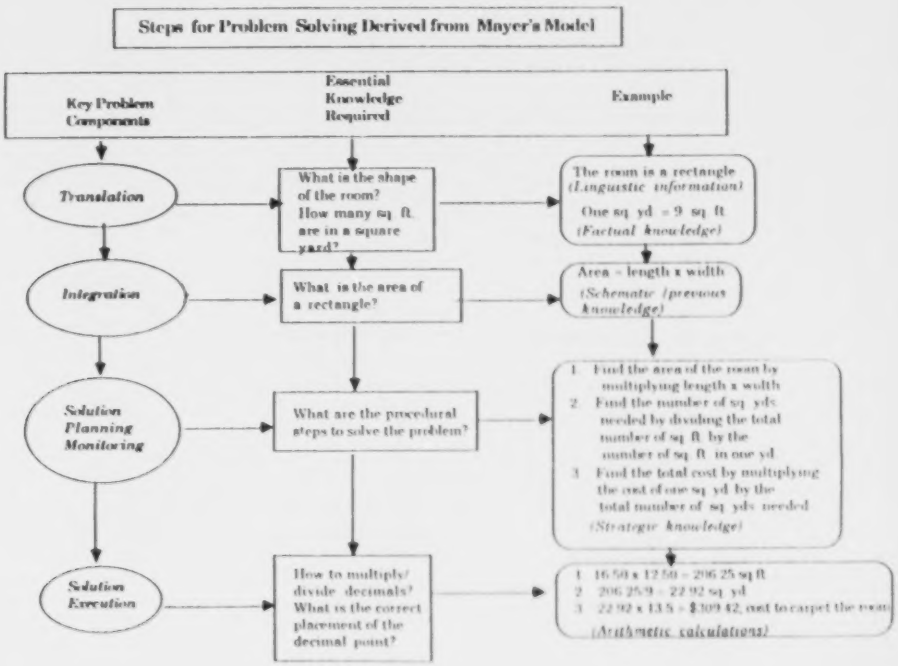
National Council of Teachers of Mathematics (1989). Standards. Reston, Va.: National Council of Teachers of Mathematics.

National Assessment of Educational Progress. (1988a). Mathematics assessment: discrete mathematics, data organization and interpretation, measurement, number, & operations. Mathematics Teacher, 81, 241-249.

Nicholls, J., Cobb, P., Wood, T., Yackel, E., & Patashnick, M. (1990). Dimensions of success in mathematics: Individual classroom differences. Journal for Research Mathematics Education.

Fig. 1
A strategy for solving mathematics problems

Problem: How much does it cost to carpet a room 16.50 ft. long and 12.50 ft. wide? The carpet cost per square yard is $13.50.


Steps for Problem Solving Derived from Mayer's Model
Key Problem Components
Essential Knowledge Required
Example
Translation
What is the shape of the room? How many sq ft are in a square yard?
The room is a rectangle (Linguistic information)
One sq yd = 9 sq ft (Factual knowledge)
Integration
What is the area of a rectangle?
Area = length x width (Schematic (previous knowledge)
Solution Planning Monitoring
What are the procedural steps to solve the problem?
1. Find the area of the room by multiplying length x width
2. Find the number of sq yds needed by dividing the total number of sq ft by the number of sq ft in one yd
3. Find the total cost by multiplying the cost of one sq yd by the total number of sq yds needed
(Strategic knowledge)
Solution Execution
How to multiply/ divide decimals? What is the correct placement of the decimal point?
1. 16.50 x 12.50 = 206.25 sq ft
2. 206.25/9 = 22.92 sq yd
3. 22.92 x 13.5 = $309.42, cost to carpet the room
(Arithmetic calculations)

HELPING BEHAVIOR AS AN INDICATION OF COGNITIVE MODIFIABILITY WITH CHILDREN IN SPECIAL KINDERGARTEN

G.M. van der Aalsvoort, & M. de Groot, Department of Special Education, University of Nijmegen, Nijmegen, The Netherlands

Introduction

Cognitive modifiability can be studied when examining behavior of the child that performs tasks designed to elicit learning potential. The child is offered by an adult in a standardized prompting procedure. The performance of the child informs us on the effect of help expressed in the gainscore but also on the way in which the child profits from help (Lidz & Thomas, 1987; Resing, 1990; van der Aalsvoort & Ruijssenaars, 1992). We state that this more dynamic assessment is preferable to a more static assessment of cognitive development since it informs us on the way in which actual performance in the classroom may be increased, ecological validity of cognitive modifiability being our final goal (Campione & Brown, 1987, 1990). Since children showing on-task behavior when asked to perform, are considered to be cognitively active and more easily elicit help of the adult in the classroom than children showing poor on-task behavior (Wagenaar & Scholte, 1988), we try to answer two questions in our study: Can on-task behavior with children, elicited when performing a Learning Potential Task, also be elicited by prompting in a classroom intervention? Does increased on-task behavior refer to increased cognitive modifiability?

Method, Design and Procedure

In our study 6 children aged 5 to 7 years from Special Kindergarten joined in a teaching experiment within a pretest-posttest-control group design (Cook & Campbell, 1979). These children were selected by their teacher because of their poor on-task behavior according to the ratings on the Wagenaar & Scholte questionnaire. After assessing the amount of on-task behavior in the classroom 6 times in a three week period and the amount of prompting when performing a classification task, a learning potential task (Hamers & Ruijssenaars, 1984)) in the pretest phase, two groups were formed of three children each. The children in the experimental group needed the most prompting compared with the children in the control group. The amount of on-task behavior of the children in the experimental and the control group was matched. The intervention consisted of performing

developmental tasks with the teacher in the classroom. The training was carried oud three times a week, for four weeks. The tasks consisted of constructing material, putting beads on a string in a certain order, and puzzles. The children of the experimental group performed the tasks on their own level but all of them puzzled etc. at the same time. The teacher used the standardized way of prompting as used in the learning potential task. The control group received no treatment. Two training sessions were videotaped. The assessment of the learning potential task was also videotaped in the pre and posttest. In the posttest phase the classification task was administered again with both the children of the experimental and the control group. The amount of on-task behavior of all participating children was assessed again 6 times in a three-week period after the intervention.

Results

The amount of on-task behavior of the pre- and posttest of the experimental group and the control group was compared and statistically analyzed by TIDA (Oud, 1986). On-task behavior of *both* groups increased. The difference between the groups was not statistically significant. Helping behavior refering to prompting of the adult and asking for help of the child during the assessment of the classification task and during the intervention sessions, of both the children of the experimental group and the adult was analyzed. qualitively. The ranking of frequency of both the verbal and non-verbal behavior of the children and the adult correlated significantly according to Pearson Rankcorrelation (Ferguson, 1971).

Conclusion and Discussion

The results of this study indicate that on-task behavior can be elicited in the classroom by prompting the child when he is asked to perform developmental tasks. The study did not show however that the intervention caused the on-task behavior since both the experimental and the control group showed increased on-task behavior. Our second question if increased on-task behavior refers to increased cognitive modifiability can be answered positively in the sense that social mediation has improved. This is showed by an increase of ranking frequency of child and adult behavior when frequency of pre- and post helping behavior are compared. Lidz & Thomas (1987) mention the possibility of improved matching as an indication of occurring cognitive modifiability.
Adequate helping behavior may be more informative on *how* cognitive modifiability is elicited with children in the classroom, being a prerequisite to cognitive modifiability.

References

Aalsvoort, G.M. van der, & Ruijssenaars, A.J.J.M. (1992). Kwalitatieve analyse van hulpgedrag als teken van cognitieve veranderbaarheid bij het oplossen van classificatietaken. Tijdschrift voor Orthopedagogiek, 32, 68-81.

Campione, J.C., & Brown, A.L. (1987). Linking Dynamic Assessment with School achievement. In C.S. Lidz (Ed.) Dynamic Assessment, an interactional approach to evaluating Learning Potential. New York: Guilford Press.

Campione, J.C., & Brown, A.L. (1990). Guided Learning and Transfer: implications for approaches to assessment. In N. Frederiksen, R. Glaser, A. Lesgold & M.G. Shafto. Diagnostic Monitoring of Skill and Knowledge Acquisition. New Jersey: Lawrence Erlbaum Ass.

Cook, T.D., & Campbell, D.T. (1979). Quasi-Experimentation. Chicago: Rand McNally.

Ferguson, G.A. (1971). Statistical Analysis in Psychology and Education. London: McGraw Hill.

Hamers, J.H.M., & Ruijssenaars, A.J.J.M. (1984). Leergeschiktheid en leertests. Een leertestonderzoek bij eerste klassers in het gewoon lager onderwijs. Lisse: Swets & Zeitlinger.

Lidz, C., & Thomas, C. (1987). The preschool learning assessment device extension of static approach. In C.S. Lidz (Ed.) Dynamic Assessment. an interactional approach to evaluating learning potential. New York: Guilford Press.

Oud, J.H.L. (1986). *TIDA:* Kunst Keyword-version. Interne publicatie, Katholieke Universiteit, Nijmegen.

Resing, W.C.M. (1990). Intelligentie en leerpotentieel. Lisse: Swets & Zeitlinger.

Wagenaar, E., & Scholte, E. (1988). Selectieve beoordeling van kleuters van verschillende etniciteit en sexe: een verklaring voor schooluitval. In J.C. Wolf & J.J. Hox. Kwaliteit van onderwijs in het geding. Amsterdam: G.P.I. publicaties.

PROPORTIONAL REASONING IN CHEMISTRY: EFFECTS OF TASK VARIABLES AND LEVEL OF INSTRUCTION

A. Sanz, J.I. Pozo, & M.P. Perez-Echeverria, Faculty of Psychology, Autonomous University of Madrid, Spain

Proportional reasoning is included across the world as a basic content in the curriculum of primary and secondary school. Usually, mathematics teachers spend a lot of time trying their students learn to solve proportional problems in the hope of this learning will be transfered to other areas with specific content (Physics, Chemistry, Geography and also everyday life). Despite these educational efforts and the considerable amount of psychological studies on problems of proportion and ratio, there is not still evidence about the nature of proportional strategies used by students with different level of instruction. Are proportional rules general abilities learnt with independence of specific content and context to which they are applied? Or are they rather specific content abilities that must be learnt in each domain in which could be useful? Both positions have been defended in psychological research.

According to piagetian theory of formal operational thinking (e.g., Inhelder & Piaget, 1955; see also Shayer & Adey, 1981) proportional reasoning is one of the eight formal operational schemata achieved and mastered in adolescence. Therefore, proportional reasoning would be a part of a general cognitive structure developed during adolescence and it would be mastered without relevant influence of subjects and task variables.

On the other hand, some authors (see Tourniaire & Pulos, 1985, for a review) have suggested that performance in proportional reasoning tasks is influenced by a number of variables. These variables would include item computational difficulty (e.g., Karplus, Pulos & Stage, 1983), task format and content (e.g., Nesher & Sukenik, 1991) and cognitive development of the subjects that solve the tasks (Noelting, 1981). However Tourniaire and Pulos (1985) suggest the influence of instruction on proportional reasoning, there is not enough evidence showing the effects of expertise in solving proportional tasks. In these problems subjects can differ in their expertise in Mathematics, in their expertise in the specific domain in which the task is presented (for instance, Chemistry proportional laws) or in both.

Objectives

This research is aimed to analyze the effects of the above mentioned variables on strategies used by subjects with different instructional level in solving proportional tasks. We are specially interested in the relationship between content knowledge -

we choose tasks with chemical content because the fundamental chemical quantitative laws are proportional in nature- and mathematical performance.

Subjects, tasks and method

We designed eight proportional tasks composed by six items with different computational difficulty. Every task was composed of one additive item -that could be correctly solved through additive, correspondence or multiplicative strategies- two correspondence items -solved using either correspondence or multiplicative strategies- and three multiplicative items -only solvable by means of right multiplicative strategies.

The eight tasks were related to scientific concepts relevant in secondary science learning. They differ in context (scientific or everyday), scientific concept implied (density, disolution or chemical reaction) and presentation format (analogical or algebraic).

We presented these tasks to 140 subjects, divided into seven different groups. Each subject solved the eight proportional tasks in two sessions according to a counterbalanced order. The seven groups were as follows: 14- and 16- years old adolescents, college preparatory students with and without chemical training, and three groups of university students from Mathematics, Chemistry and Psychology faculties. The three groups of university students had enough training in proportional reasoning but differed in their content specific knowledge (i.e., expertise in Chemistry).

Results and conclusions

We analysed both the number of right answers and the computational strategies used by each group of subjects. The number of right (i.e., mathematical) answers were analyzed through ANOVA and Scheffé's tests and also submitted to a factor analysis with VARIMAX rotation. The strategies were analyzed qualitatively.

ANOVA showed a main effect for computational item difficulty, providing empirical evidence for predicted strategies. ANOVA showed also a main significant effect for concept implied, with disolutions and density easier than chemical reactions, and for level of instruction, with the three university groups and college preparatory students with chemical training showing a better performance than the three others groups. In general, training in proportional reasoning was more related to performance than chemical knowledge. In this way, psychologists did not differ significantly from mathematics and chemists. These result could be interpreted as consistent with a content-independent approach to proportional reasoning instruction.

The effects of content and format presentation were significant only for specific groups and items. Usually, subjects performed better in everyday tasks than in Chemistry taks, but the differences were significant only for the less able subjects. Likewise, scores were higher for analogical presentation than for algebraic format (a result partly opposite to those found recently by Nesher & Sukenik, 1991). But the differences were again significant only for the groups with lower performance. But, as these differences appeared in adolescent groups, when proportional reasoning is usually taught and, for this reason, they are relevant for educational purposes, showing the introducing proportional rules in non-formal contexts.

Factor analysis showed two main factors. The first factor loaded in disolutions and density tasks and accounted for 45 per cent of whole variance, while the second accounted for 16.7 per cent and loaded in chemical reactions tasks.
Finally, analysis of computational strategies showed that subjects with higher instructional levels employed different strategies depending of specific item demands. Subjects with expertise in proportional reasoning used proportional reasoning strategies in a flexible, task-adapted way. In general, subjects tended to select strategies according to item computational difficulty, using in each item the simpler strategy that provides a right answer. Subjets with low-training in proportional reasoning did not show this pattern of strategig knowledge. Expertise in proportional reasoning apparently implies a more flexible use of strategic knowledge. It is required more research for discovering the cues used by expert subjects to select or trigger each strategy in order to fit their performance into tasks demands.

References

Inhelder, B., & Piaget, J. (1955). De la logique de l'enfant a la logique de l'adolescent. Paris: P.U.F.

Karplus, R., Pulos, S., & Stage, E.K. (1983). Proportional reasoning of early adolescents. In: R.Lesh & M.Landau (Eds.) Acquisition of mathematics concepts and processes. New York: Academic Press.

Nesher, P., & Sukenik, M. (1991). The effect of formal representation on the learning of ratio concepts. Learning and Instruction, 1, 161-175.

Noelting, G. (1983). Qualitative and quantitative aspects in the development of proportional reasoning. In: M.P.Friedman, J.P.Das & N.O'Connor (Eds.) Intelligence and learning. New York: Plenum Press.

Shayer, M., & Adey, P. (1981). Towards a science of science teaching. London: Heinemann Educational Books.

Tourniaire, F., & Pulos, S. (1985). Proportional reasoning: a review of the literature. Educational Studies in Mathematics, 16, 181-204.

STUDENTS' CONCEPTIONS ABOUT CONSERVATION OF MATTER

M.A Gómez-Crespo, & J.I. Pozo, Faculty of Psychology, Autonomous University of Madrid, Spain

Research on science learning and teaching has followed two basic lines. The first one is the *Piagetian formal thought theory* (Inhelder & Piaget, 1955; Shayer & Adey, 1985). The second, more recent, is the study of the preconceptions and *alternative conceptions* that students have about scientific phenomena (e.g., Driver, Guesne & Tiberghien, 1985).

Following this last line, there have appeared many reports related to Chemistry. Most of them are merely descriptive, without trying to find any theoretical connection between the different conceptions and the different parts of Chemistry. However, literature offers some valuable attempts to systematize the students' ideas, although they are only centred on combustion and oxidation study. Meheut *et al.*, (1985) groupe pupils' ideas on these topics into two categories. But the most interesting attempt is that developed by Andersson (1986). He has created a general scheme for the pupils' ideas on chemical reactions. More recently, Andersson (1990) has applied this scheme to changes of state. His analysis is based on five categories:

- It is just like that;
- Displacement: new substances can appear simply because they have been displaced;
- Modification: what appears to be a new substance is the same substance as before, although in a modified form;
- Transmutation: the original substance is transformed into a completely new one;
- Chemical interaction: correct view of chemical changes.

From these five categories we could study students' ideas about chemical reactions in general, not only in oxidation and reduction. But we can wonder, beyond the study of the ideas about chemical change, to what extent these categories will be useful too to analyse students'ideas about conservation of matter after physical and chemical changes.

Objectives

We try to study in which extent five Andersson's categories are useful to study students' ideas about conservation of matter, analysing if they permit us to differentiate between distinct changes matter can suffer (solution, change of state and chemical reaction).

Method: Subjects, tasks and data-analysis

We worked with 120 students from different age and instruction level (6 groups of 20 subjects): 12-13 years old, 14-15 years old, 16-17 years old (with scientific studies), 16-17 years old (without scientific studies), Psychology students (University, last year) and Chemistry students (University, last year).

In the context of a wider research, we designed two 11-item tests about matter conservation with two different presentations (chemical and everyday life). Every item had five options of answer corresponding to five categories, adapted from those created by Andersson (1986, 1990):

- Interaction. The substances interact to form a new one;
- Displacement. Substances appear or disappear;
- Transmutation. Some substances change into others;
- Modification with identity. The substance modifies its aspect but it continues to be the same;
- Modification of amount. The substance continues to be the same but the amount in different.

In order to analyse the results we have used two criteria: number of correct answers and category used to answer. With regard to the first, we have used an analysis of variance (ANOVA) with a factorial design 6x2x3, between-groups for the first factor (group) and intra-groups for the other two (test and contents). For the second, we have used an analysis of variance (ANOVA) with a factorial design 6x2x3, between-groups for the first factor (group) and intra-groups for the other two (contents and categories).

Results and conclusions

ANOVA of the *number of correct answers* show these significant main effects:

- *Group* variable ($F=38.345$ $p<0.0001$). Score increases with age and instruction in Chemistry. However these results show a picture far from unexpected, there appeared no differences between subjects (16-17) with and without scientific studies and, also, Psychology students, which are naive in Chemistry, attained results better than expected. These results show a smaller influence of chemical expertise than that predicted from research in other scientific areas (e.g., Pozo & Carretero, 1992). Expertise makes also lesser differences than that found in other chemical contents, as proportional reasoning in chemical tasks (Sanz, Pozo & Perez-Echeverria, 1992).
- *Content* variable ($F=3.535$ $p<0.0305$). The most difficult contents are *reactions* (when task is presented in an *every day life* context) and *solutions* (when it is presented in a *chemical* context). There are difficulty differences when trying to explain conservation in different chemical problems; so, *change of state* is the easiest content for subjects. For this content, task presentation seems to be no relevant, unlike what happens with solutions and reactions.

The main effects shown by ANOVA in relation to the answer *categories* were the following:

- *Category* variable ($F=304.150$ $p<0.0001$). The most used categories to answer are interaction (40 %) and modification with identity (40 %). Transmutation category (12 %) is less used. Displacement (3 %) and modification of amount (5 %) categories are practically not used.

- *Test-category* interaction ($F=7.898$ $p<0.0001$). Interaction category is used more times for chemical context (45 %) than for every day life context (35 %).
- *Content-category* interaction ($F=107.352$ $p<0.0001$).
- *Solutions.* The most used category is modification with identity (58 %, correct category), although there is an important group using interaction category (30 %).
- *Changes of state.* Students use basically modification with identity category (60 %, correct category), although there is a little group using interaction (13 %).
- *Reactions.* The most used category is interaction (39 %, correct category), although there is a rather important use of transmutation (16 %) and modification with identity (17 %).

These results show that answers are influenced by the way the phenomenon is presented. So, if chemical terms are used, they tend to interpret it as interaction between substances to produce some new ones more often than if the presentation is every day life. This trend is found when the process includes more than one substance: solutions and reactions. Although Andersson (1986, 1990) hardly gives quantitative data, we have noticed that displacement category seems to be used less often than supposed. In the opposite point, interaction category is used to interpret chemical reactions much more often than anticipated, and it is also misused to interpret physical changes in chemical terms.

References

Andersson, B. (1986). Pupils' explanations of some aspects of chemical reactions. Science Education, 70(5), 549-563.

Andersson, B. (1990). Pupils' conceptions of matter and its transformations (age 12-16). Studies in Science Education, 18, 53-85.

Driver, R., Guesne, E., & Tiberghien, A. (1985). Children's ideas in science. Milton Keynes: Open University Press.

Inhelder, B., & Piaget, J. (1955). De la logique de l'enfant a la logique de l'adolescent. Paris: P.U.F.

Meheut, M.; Saltiel, E., & Tiberghien, A. (1985). Pupils' (11-12 year olds) conceptions of combustion. European Journal of Science Education, 7(1), 83-93.

Pozo, J.I., & Carretero (1992). Causal theories, reasoning strategies and cognitive conflicts in Newtonian Mechanics. In: A.Demetriou; M.Shayer & A.Efklides (eds.) Modern theories of cognitive development go to school. London: Routledge and Kegan Paul.

Sanz, A., Pozo, J.I., & Perez-Echeverria, M.P. (1992). Proportional reasoning in Chemistry: effects of task variables and level of instruction. Paper presented at European Conference of Educational Research, ECER, Enschede, The Netherlands, june 1992.

Shayer, M., & Adey, P. (1981). Towards a science of science teaching. London: Heinemann Educational Books.

SCIENCE ACTIVITIES AS A MODEL OF COOPERATIVE SOCIAL BEHAVIOUR FOR ADOLESCENTS WITH CONDUCT DISORDER

B. R. Zinn, & D. I. Kohn, Rothberg School for Overseas Students, The Hebrew University of Jerusalem, Israel

Introduction

Can science activities provide a positive personal experience for adolescents with conduct disorders? Their behaviour patterns persistently violate the the rights of others and major societal norms. They lack limits regarding themselves, their relations with others and learning tasks. Since they recognise no rules, there is no evident cause and effect governing their their surroundings which leads them to perceive the world as ruled by magic powers. These pupils are characterised by their distrust and the heavy reliance they place on their senses as their only means of interpreting their environment.

The nature of science lends itself to cooperating and communicating with others as well as seeking less subjective ways of explaining the environment. It is postulated that by using hands-on activities pupils will not only learn about causality and science process skills, but they will also be able to confirm rules or principles by testing and retesting for themselves. Different but interrelated science activities can provide a positive personal experience which should help improve attitudes to scholastic tasks, learning environment and self-image.

The purpose of this study is to describe the rationale and activities of a recently developed science project for special education pupils. Preliminary findings which examine the feasibility of implementing these activities are reported.

Rationale

One of the characteristics of these pupils in a school setting is that their cognitive performance is noticeably lowered, while their potential, as measured by standard tests, is within the bounds of normality. They are so involved with their own inner conflicts that they have a short attention span for academic tasks as they are easily disrupted by any external or internal stimulus.

Science activities are suggested for the following reasons:

- The content of natural science is neutral and less emotional than in social science.
- Causality in natural science is more definite and easier to explain than in social science.
- Concept formation is based on natural and concrete phenomena.
- Concepts, rules or principles can be confirmed by student experiment.

- Duration of experiments can be short and criteria for success in many cases can originate from the activity itself.

Development of activities

The program consists of a series of modular activities. The activities are short so that pupils can experience immediate positive feedback. As a first step, the following eight science process skills were identified as the core aims of the project :

- making observations;
- finding similarities and differences;
- collecting and organising data;
- exploring the need for standardised measurements;
- making predictions;
- performing experiments;
- identifiying cause and effect;
- drawing conclusions from experimental data;

The second step was to identify content areas which lent themselves to hands-on activities. Discrepant events were sought as they have a motiational effect on the learner who faces a situation which is in conflict with what he or she expects. These two steps were combined into a matrix od specific content areas matched with the eight skills. The third step involved the development of worksheets, which were written with precise instructions and concise explanations with much graphic support. Closed structured questions were used so that pupils had to chose between alternatives.

The activities deal with exploring the senseswhile showing their limitations, performing chemical reactions to make toothpaste and soap, investigating common household acids and bases by using red cabbage as an indicator, making predictions with soap bubbles and magnets.

Practical considerations dictated a simple pre and post test design with an experimental and control group. Even though the entire school participated, numbers are small. We are dealing with a special education school whose only intake is pupils with conduct disorders.

Findings

The pre test showed interesting misconceptions such as :

- if you eat too many sweets, your teeth will get fat
- soap bubblesare flat
- soap bubbles take on the shape of the soap
- temperature is measured with a ruler
- to remove a greasy stain from a shirt, you simply throw out the shirt

These misconceptions were cleared up by the program. Causality as well as science process skills were reinforced.

On an affective level there was considerable improvement in their attitude towards the learning environment which was reflected in longer attention spans, following instructions, completing tasks, cooperating with each other and a non-destructive approach towards the experimental materials. Pupils suggested and performed futher simple experiments stimulated by their positive experiences during the activities. They were motivated to come each week and looked forward to each

session. They reported the science activities as being interesting and fun in the school newspaper. They also related their experiences to pupils in the control group, who clamoured to be included in the next run.

Conclusion

These preliminary findings indicate that pupils with conduct disorders can learn about causality. Their self confidence can be improved by achieving success through their own activites. This model of using hands-on science activities with adolescents with conduct disorders merits futher investigation.

AGE RELATED AND INDIVIDUAL DIFFERENCES IN THE SOLUTION OF VERBAL ANALOGIES

Wilma C.M. Resing, & Nicole Berger, University of Leiden, Department of Developmental and Educational Psychology, The Netherlands

Abstract

Previous research suggests large individual and age differences in the frequencies at which children use word associations instead of analogical reasoning when solving verbal analogy problems. Young and less able verbal analogy solvers seem to be more associative reasoners and show less inference and application in the solution process than older and more capable solvers do. They also do not fully encode the more complex stimuli.

The emphasis in this study lies on comparing process outcomes in successful and less successful verbal reasoners within three age groups, by contrasting their results on two versions of the same verbal analogies test: a generation task (open version) and a multiple choice task (forced choice version).

Subjects were 225 2-, 4-, and 6-grade elementary school children. Two versions of a verbal analogy task were developed with a set of systematically chosen answering alternatives. Based on a pilot study associative distractors were chosen in addition to the correct alternative. Based on both test registrations, five different components can be estimated: P (generation), P (forced choice), P recognition, P distraction and P guessing. By obtaining estimates for each of these components, the contribution of each component to age and skill differences can be established. The children got both tasks in fixed order of presentation.

The results of the study will be presented at the conference.

Introduction to the problem

The ability to reason by induction has a central place in cognitive development (e.g., Bruner, 1973, Goldman & Pellegrino, 1984). Inductive reasoning is also important in investigations concerning human intelligence. In several intelligence theories (Spearman, Vernon, Sternberg, modern radex theories) inductive reasoning takes a central position. Most intelligence tests contain one or more of these tasks such as matrices, analogies, series, inclusion, etc.

Verbal analogies mostly have the form of A:B = C:??? (open version) or A:B =C: (D1 - Dx) (forced choice version). In Sternberg's (1977, 1985) theory on analogy solution, the terms A and B of the analogy have to be encoded and the relationship between A and B has to be inferred. Then C has to be encoded and the information about A and B has to be applied to the C term. In addition the

correspondence between the A-B and C-D pairs of the analogy has to be established (mapping).
Previous research suggests that young children (8 years) do not spontaneously use the mapping component in the solution process and that these children do not fully encode the more complex stimuli. They solve the problems less efficient than older children do (Sternberg & Nigro, 1980; Goldman & Pellegrino, 1984). Achenbach (1969, 1970) investigated the responses children gave on verbal analogy items. He found large individual and age differences in the frequencies in which children used word associations (on the C-term of the analogy) instead of analogical reasoning when solving the problems. Sternberg and Nigro (1980) conclude that association affects but does not wholly control analogy solution. Young and less capable verbal analogy solvers seem to be more associative reasoners and show less inference and application than older and more capable solvers do. In this study we wanted to replicate these findings of Goldman, Pellegrino, Parseghian and Sallis (1982) for children from second, fourth and sixth grade. Instead of verbal analogy items from existing intelligence batteries (see Goldman et al, 1982) which sometimes have verbal associations as the correct answer, we used a systematically constructed verbal analogy test with among the answers in the alternative set strong and less strong verbal associations with the C-term of the analogy that are not the correct answers.
The emphasis in our study lies on comparing process outcomes in successful and less successful verbal reasoners within three age groups, by contrasting their results on two version of the same verbal analogies test: a generation task (open version) and a multiple choice task (a forced choice version).

Method

Subjects

The subjects were 75 second-grade, 75 fourth-grade, and 75 sixth-grade children from 3 elementary schools in the Netherlands.

Tasks

Two versions of a verbal analogy task were developed, with 30 items each: an open version and a version with the same items but with a set of five answer alternatives. In a pilot study children from grade 2, 4 and 6 had to describe their associations with the C-terms of the analogies. Based on this pilot study distractor alternatives were chosen in addition to the correct alternative and other ones. The items represent different types of analogy relations, such as function, class membership, property and part-whole relations.

Measures

The items from the generation task can be solved correctly when the individual infers the correct relation between the A and B term of the analogy and then applies that relation correctly on the C-term to find an answer. The level of correct performance in the generation task is given by the probability of finding both the correct inference and the correct application. The presence of alternatives in the forced choice task can facilitate task performance. Facilitating elements are guessing and recognition processes. However, the presence of the alternative set also can items make more difficult to solve, especially when one or more of the

incorrect alternatives is a strong free association with the C-term of the analogy. The level of performance in the forced choice task is given by the probability to chose the correct answer in the generation task, corrected for guessing, recognition (P_r) and distraction (P_d). The following four probabilities can be estimated: $Pcorr_{gen}$, the probability that the correct answer is chosen in the generation task; $Pcorr_{fc}$, the probability that the correct answer is chosen in the forced choice task, P_r, the probability that for items with incorrect generation response a correct forced choice response is given; P_d, the probability that for items with correct generation response an incorrect forced choice response is given. By obtaining estimates for each of these components, the contribution of each component to age and skill differences can be established.

Procedure

The tasks were presented over the course of 2 weeks. All the children firstly got the generation task and secondly the forced choice task. After an intensive classical instruction, the children completed the test items. For the youngest group of children the experimenter read aloud the words of the analogies before the children filled in their answers.

Results

The results are not available at this moment and will presented at the conference. Mean values for the different process components will be estimated for age an skill (high- and low-scorers) groups. Simple and multiple regression analyses will be performed for the prediction of the multiple choice performance.
We expect that young and less skilled children will show more distractor interference than older and more skilled children. In addition we expect that the estimated values for distractor interference in this study will be slightly higher than those in the study by Goldman et al., at least for the youngest children.

COOPERATIVE LEARNING AND TEXT COMPREHENSION: A SIMULATION OF REAL CLASSROOM PRACTICE

Josefa Guillén, Dolores Rodriguez, David Saldaña, Inés Sola, Developmental and Educational Psychology, University of Seville, ESP

There is a widespread belief in Spanish schools that cooperative learning is, as a rule, better than individual or competitive learning. Much research has been carried out on this matter (Slavin et al, 1.985) that points to certain factors which seem to have great influence on results of the use of these methods on academic achievement.

In our immediate context, however, these findings do not seem to have brought about much change in existing practices. Cooperative learning is, in general, adopted by well–meaning teachers in all levels of education, without an adequate organization of task or reward structure, and with no regard for its effects on individual learning.

This experiment intends to simulate usual classroom practices of cooperative learning. For this purpose, the task of producing an outline from a text has been chosen. This task not only requires intensive use and handling of information extracted from the text, but, presumably, also has favourable effects on reading comprehension. It is normally assumed that class work of this sort is of superior quality when carried out as a group task than individually, mainly because of free interactions among students. Following this reasoning, the hypothesis that guide this study are:

- The outlines produced as group tasks will be of superior quality than those produced individually.
- Collective work on the outline favours reading comprehension more than individual work.

Method

Subjects

The study was carried out with 52 Spanish high school students (20 males and 32 females, with an average age of 15:5).

Instruments

Two texts, similar in difficulty, structure and number of words, were adapted from First Year History and Geography high school books for use in this investigation. Two 6–question comprehension tests were designed specifically for the investigation (title definition – dimension 1 –, two low level questions – dimension

2 –, two slightly higher level questions – implying hierarchichal and classification strategies – dimension 3 – and a last question requiring extension of acquired information to real–life situations –dimension 4).

Design

A pre–postest experimental design with control group was used. All subjects first produced an individual outline of text1, after which they responded to the corresponding comprehension test.In the next session, carried out on the following day, the students were randomly assigned to the experimental (team–work) or control (individual) condition. Subjects in the experimental group worked on the outline of text2 in various teams that had to produce their own outlines, after which they were asked to respond individually to the corresponding comprehension test. Students in the control group were asked to complete the outline and the test individually.

Results

Pretest measures for both experimental and control groups were compared with t–Student tests, and ANOVA tests were carried out on postest scores.
Results allow us to conclude that there was no statistically significant difference on any dimension in reading comprehension nor on the outline measures. The only exception was dimension 2 of the reading comprehension measure, in which there were significant differences in favour of the control condition ($p<.05$).

Discussion

Hypotheses 1 has not, in view of these results, been confirmed, i.e., cooperative work on the outline has not produced higher quality than individual work. Non–confirmation of hypotheses 2 was, therefore, to be expected: since cooperative work proves no better than individual work, there is no reason for reading comprehension to have improved in those students that took part in it. The fact that measures on dimension 2 have even worsened would suggest the possibility that not all tasks benefit from cooperative work, but should, on the contrary, be carried out individually. This conclusion should, however, be confirmed through further research.
There are various factors that could explain these results, namely:
Lack of control over classroom structure:

- task structure;
- individual accountability;
- reward structure.

Lack of organization of interaction among:

- teacher–students;
- students–students (peers).

The structure followed in our study is similar to the one commonly found in High Schools in our cultural context. It follows that, unless there is a stricter planning of this type of work, nothing much is being gained through its use. And since what, ultimately, seems to be the determining factor in promoting academic and cognitive gains is the actual interaction among students and with teachers (Coll, 1984; Mora, 1991), it is in this direction that futher research is required.

405

BEST COPY AVAILABLE

References

Coll, C. (1984). Estructura Grupal, Interaccion entre Alumnos y Aprendizaje Escolar. [Group Structure, Student Interaction and Learning at School] Infancia y Aprendizaje, 27/28, 119–138.

Mora, J. (1991). Evaluacion de un Programa de Enriquecimiento Cognitivo: Comprender y Transformar.[Evaluation of a cognitive enrichment program: Understanding and Changing]. Sevilla: Unpublished Doctoral Dissertation.

Slavin, R.E. (1985). An Introduction to Cooperative Learning Research. En R.E. Slavin et al. Learning to Cooperate, Cooperating to Learn. N.Y.: Plenum Press.

406

MOTIVATION AND SELF REGULATION

E. De Corte, University of Leuven, Leuven, Belgium

Two years ago the Dutch Foundation for Educational research (SV0) funded a strategic long term research programme entitled "Motivation and self-regulation". This programme aims to integrate three until now relatively independent research traditions: being motivational research, research on self-regulation and research on teacher influences on the development of student autonomy. The research hypotheses are:

- The motivational orientation is co-determined by successes students have with learning activities;
- The motivational orientation is co-determined by opportunities students have to employ learning activities on their own and to take initiatives;
- School career is co-determined by motivational orientation.

An important impetus for the project was a study done by Helmke and Weinert from Munich. The general hypothesis resulting from their study and being the central focus in our project is that the direct teaching model (maximizing task relevant activities high performance expectations) is good for performance. self-concept of ability and fear but that it leads to lower motivation for mathematics and a more negative attitude towards school. This implies that autonomous learning activities by students themselves should perhaps get more attention from teachers and researchers.

The project consists of three integrated projects:

- Motivation. action control and performance;
- Self-regulation skills and the choice of learning goals and learning activities and requests for help in computer aided instruction;
- The influence of the quality of instruction on the motivational orientation and the school-career of students.

One major aim of the research programme is to develop a set of instruments that can reliably assess a student's ability to regulate its own learning. both in a cognitive and a motivational sense. The first project adapted Kuhl's Action-control scale for secondary education and investigated its effects on student s learning intention and their effort expenditure. The second project studied self-regulation skills of students in choosing learning goals and learning activities as well as requests for help. The third project studied the influence of teacher behaviour with respect to aspects of the direct teaching model and activation of student activities on motivation.

At this stage the three research projects finished instrument construction and the various instruments are currently being used in a longitudinal field study with

students from secondary education. Instruments from the three projects are related to each other in a three year longitudinal study with about 1000 students from 6 secondary schools. At present the data from the first year students have been collected. Relations with several school-career variables will be studied: (Real and virtual) choices of school-subjects choice of activities and goals and learning performance.
The symposium will consist of five parts. In the first part the rationale behind the research programme will be presented. In the second part the first emanating results from the three research projects will be reported. In the third part the design of the longitudinal study will be presented as well as an attempt at integration. The fourth part consists of a paper presented by Andreas Helmke from the Max Planck Institute (München) about the results from the Munich longitudinal study mentioned above. Finally Gabi Salomon from the University of Arizona will act as the discussant of the symposium. There will be time for the audience to participate in the discussion.

MOTIVATION AND SELF-REGULATION: BACKGROUNDS AND THEORIES

M. Boekaerts, University of Leiden, Leiden, & P.R.J. Simons, University of Nijmegen, Nijmegen, The Netherlands

In this paper the rationale behind the programme will be presented and the relations between the projects will be explained. An overview of the research questions and hypotheses underlying the programme will be discussed in the light of prior research done at the Max Planck Institute in Munich and of theories about motivational orientations and self-regulation. One central idea of the project is that motivational theories and theories of metacognition and self-regulation have much in common and that this common ground should be investigated Metacognition should not only refer to knowledge and regulation of cognitive aspects of learning and knowing, but also to affective and motivational aspects. Kuhl's theory of action control provides possibilities for this integration. Another important idea refers to the supposed important role of student autonomy. It is expected that if and when students are able to get opportunities for self regulation of their learning activities both their motivational orientations and their self-regulation skills will develop.

METAMOTIVATION AND EFFORT

M. Boekaerts, & R. Otten, Centre for study of Education and Instruction, The Netherlands

Teachers consider various causes when they observe lack of effort in their students, for example, students are too occupied with other things, they do not feel at home in class, or they find the subject-matter too complex or boring. These explanations have in common that they use a student's motivation as the explaining variable.

Since Kuhl's Action-control theory, we know that such explanations are only half of the story. He explained that low effort does not necessarily mean that a student is not motivated. Sometimes, a behavioural intention exists, but it falls short of willpower. Carrying out an intention requires action-control, more specifically the skill to ".... protect a current intention from being replaced... before the intended action is completed (Kuhl 1984 p. 102). In an educational context it is important to distinguish between motivation problems persé and problems stemming from failing action-control skills.
In our study we investigated the relationship between action-control on the one hand and the student s motivated behaviour on the other. Kuhl's action-control questionnaire was administered to students from the first form of secondary education. Motivated behaviour was measured with the On-line Motivation questionnaire: students' task specific cognitions, their emotions, as well as their intended and actual effort were measured vis-a-vis four different school-subjects: viz., French, history, Native language, and mathematics.

THE INFLUENCE OF TEACHERS' SENSE OF EFFICACY AND MOTIVATION ON THEIR INSTRUCTIONAL BEHAVIOUR

J. van Amelsvoort, & Th. Bergen, University of Nijmegen, Nijmegen, The Netherlands

Regarding the discussions on this topic, two aspects of quality of instruction 'direct teaching' and teaching which supports selfregulation of students. seem difficult to combine (Weinert, Shrader, & Helmke). Although direct teaching promotes achievement, it decreases the chance that students will develop autonomy and strategies of self regulation. For the instructional process, the development of autonomy and strategies of self regulation means that teachers should stress the activation of students' learning processes instead of controlling them extensively. This paper reports research on 69 secondary education teachers, in which we examined the differences in teachers' instructional behaviour with respect to activation and control of students under the influence of their sense of teaching efficacy and their motivation to teach. Measures of teachers' sense of teaching efficacy and their motivation to teach are adopted from the studies of Den Hertog and Van Opdorp. For the measurement of activation and control we developed the Questionnaire on Instructional Behaviour (QIB) which focuses on teachers classroom behaviour from an instructional point of view This questionnaire is a perceptual measure and can be used with students as well as teachers. Filled in by teachers themselves we obtain information about the way teachers view their own instructional behaviour. The results show that teachers with a higher sense of teaching efficacy view their instructional behaviour as more clear more controlling and leaving more responsibility to students. Motivation does not correlate significantly with instructional behaviour. These results suggest that teachers with a higher sense of teaching efficacy are more able to combine aspects of controlling and activation than teachers with a lower sense of efficacy.

COMPUTERBASED ASSESSMENT OF SELFREGULATION ACTIVITIES DURING STUDYING TEXT

F.P.C.M. de Jong, C. Kluvers, & P.R.J. Simons, University of Nijmegen, Nijmegen, The Netherlands

Assessment of selfregulation strategies is mostly done by means of questionnaires, error detection tasks or by use of the thinking aloud method. Each instrument has its own problems. Data from questionnaires only reveal what students think they should do in certain hypothetic situations. A problem of error detection tasks is that incongruities in texts are so natural, that readers are not really focused on them. Thinking aloud does map on line regulation strategies, but the laborious analysis of the protocols makes it impossible to get results in a short time. This, for instance, is needed to examine the effects of trainingprograms on the use of regulation activities or to diagnose deficiencies in students regulation behaviour to tune a trainingprogram to the needs of the student. Our efforts are focused on the development of a computerprogram to diagnose selfregulation behaviour. The program, CASA (Computerbased Assessment of Selfregulation Activities) diagnoses nine regulation activities: choosing adequate learning goals, choosing adequate learning activities, planning learning activities considering learning goals using prior knowledge, reacting adequate (action oriented) on comprehension problems, getting help at the right moments, monitoring the learning process checking the learning process and using constructive learning activities. The data we will discuss result from a study in which 50 students of Secondary Education participated. Students answered the questions of the Inventory Learning Styles (ILS), of the Inventory Regulation Strategies (IRS), made performance tests (as pre- and posttest) related to the text to be studied in CASA and accomplished the CASA diagnosis. The data of the questionnaires are correlated with the CASA-data in order to validate the CASA-instrument.

MOTIVATION, SELF-REGULATION AND SCHOOL-CAREER VARIABLES IN A LONGITUDINAL PERSPECTIVE

Th. Bergen, University of Nijmegen, Nijmegen, The Netherlands

The next step within the research programme 'Motivation and self-regulation' will focus on the development of motivation and self-regulation of students during the first phase of secondary education (age 12-16 years). At the symposium the research designs of the longitudinal studies of the programme will be presented. The first research aim of the longitudinal phase of the programme is to examine the influence of motivation and self-regulation on the schoolcareer variables in the schoolsubjects Dutch, English Mathematics and History. The second aim is to replicate the results of the study of Helmke for the Dutch situation by studying the relationship between teaching behaviour motivation and performance. The third aim is to figure out the conceptual similarities or dissimilarities between motivation and self-regulation in a task specific situation.

MOTIVATION AND PERFORMANCE IN A LONGITUDINAL PERSPECTIVE (PROVISIONAL TITLE)

A. Helmke, Max Planck Institute, Germany

In a two year long period the relations between teaching behaviour, motivation and performance were studied. At the end of the two year certain teaching behaviours (fitting in the direct teaching model) correlated with performance on a mathematics achievement test self-concept of mathematical ability and fear of mathematics. however, there were negative correlations with motivation for mathematics and attitude towards school. What was good for performance, thus .seemed to be bad for motivation. Implications for further research will be discussed.

411

THE ROLE OF DOMAIN–SPECIFIC KNOWLEDGE AND EXPERTISE IN LEARNING AND INSTRUCTION: AN INTRODUCTION

Filip Dochy, Centre for Educational Technology and Innovation, Open University, Heerlen, The Netherlands, & Erik De Corte, Centre for Instructional Psychology and Technology, University of Leuven, Leuven, Belgium

After the golden age of general heuristics, it has become more and more clear that the specific content of the study is sometimes strictly related to the applicability of certain heuristics. How people apply their knowledge and a further understanding of the process seems today only possible when at least taking the domain–specificity into account.

Domain–specificity is the empirically based view that learning (a) requires prior knowledge, (b) is unlikely to transfer from one domain to another without explicit transfer–inducing instruction, and (c) is unlikely to be learned from general learning of thinking instruction.

This symposium starts from the central statement in several theories related to the role of expertise in learning, such as schema theories. This statement says that the schemata or knowledge structures a person already possesses are a principal determiner of what will be learned. Recent research has showed that domain–specific expertise plays a important role in this (Weinert et al., 1990; Dochy, 1992). The symposium will focus at the fact that domain–specific knowledge plays a much greater role than what was assumed earlier. The papers are related to methods for measuring domain–specific knowledge, the role of domain–specific prior–knowledge during the learning process and its power for explaining variance in study results.

References

Weinert, F., Schrader, F.–W., & Helmke, A. (1990). Quality of instruction and achievement outcomes. Learning & Instruction, 13, 8, 895–914.

Dochy, F.J.R.C. (1992). Assessment of prior knowledge as a determinant for future learning. Utrecht/London: Lemma/Jessica Kingsley Publishers.

412

SUBJECT MATTER KNOWLEDGE AND TEACHING PERFORMANCE

Neville Bennett, University of Exeter, School of Education, Great Britain

In recent research on thinking and teacher thinking the emphasis has shifted from what teachers do to the knowledge the teachers hold, how they organise that knowledge and how various knowledge bases inform their teaching. Shulman (1987) for example has identified areas of teacher knowledge each with hypothesised relationships to teacher action. Research undertaken from this perspective has emphasised the importance of subject matter knowledge, both substantive and syntactic, in enabling teachers to pose questions, to explain concepts, to select tasks, evaluate student understandings and make curricular choices (cf McDiarmid, Ball and Anderson, 1989). However these studies have been undertaken on small samples of American student teachers, largely at the secondary level.

The study to be reported emanates from a larger research programme on the training of primary school teachers, focusing on the development of teaching competences. The question addressed in this paper is the relationship between students subject knowledge and their teaching performances in the classroom; specifically, do students teach their specialist subject at higher levels of competence than non–specialists?

A sample of 12 students, four from each of three specialist courses – mathematics, science and music, were selected for study. Each was observed for 4 days, for a minimum of 10 lessons. For each lesson the following data were collected: the student completed a lesson planning protocol detailing their intentions for the lesson, the tasks chosen, how they were to be presented, modes of assessment and classroom organisation. The lesson was audio recorded and observed by a trained observer, who then carried out a post–lesson interview with the student on their reflections and evaluations of the lesson.

Transcripts and interview data have been analyzed using a category system based on the distinction between the intellectual and logistical aspects of teaching (Kennedy, 1991). Analyses will be presented contrasting lessons taught by specialist and non–specialists, together with case studies chosen to highlight the impact of subject matter knowledge on the quality of teaching performances.

References

Kennedy M. (1991). 'An agenda for research on teacher learning'. NCRTL, Michigan State University.

Shulman L.S. (1987). 'Knowledge and teaching: foundations of the new reform'. Harvard Educational Review, 57, 1–22.

McDiarmid, G.L., Ball, D.L., & Anderson, C.W. (1989). Why staying one chapter ahead doesn't really work: subject specific pedagogy. In Reynolds, M.C. (Ed). Knowledge Base for the beginning Teacher. New York: ACTE Pergamon.

413

THE ROLE OF DOMAIN–SPECIFIC KNOWLEDGE IN THE DOMAIN OF DESIGNING

Henri H.C.M. Christiaans, School of Industrial Design Engineering, Delft University of Technology, Delft, The Netherlands

In this paper we focus on the domain of designing and especially to the question what role the knowledge of a designer is playing during the process. With the term knowledge we refer to " an individual's personal stock of information, skills, experiences, beliefs and memories". The concept of information refers to the input–side, before it is processed by a person.
There are several reasons motivating this surge of attention. First, a new paradigm in design methodology is the idea that for a fruitful design theory we badly need a description of the design process in terms of cognitive and computational models. Second, research in cognitive psychology during the last decades has made clear that knowledge is of major relevance in understanding problem solving. Third, the work on knowledge–based problem solving supports the view that expert problem solving depends primarily on having the appropriate domain–specific knowledge and not on any unusual intellectual abilities (Anderson, 1987; Elio & Scharf, 1990). With domain–specific knowledge we mean knowledge about a specific field of study, her referring to industrial design engineering.
The theoretical background of knowledge acquisition and learning is primarily based on studies into the process of solving well–structured problems, like problems in physics and mathematics. Studies of ill–structured problems in domains like design engineering are less frequently mentioned. Reason for this is probably the difficulty encountered in analyzing solution processes for this kind of problems, with their typically heuristic nature. Since it is difficult to give precise descriptions of complex solution processes, rules and methods for the educational practice are not available. What education needs is knowledge that can be explicitly taught, rather than the often very implicit way in which knowledge about complex solution processes must be mastered nowadays during individual or group exercises.
Teaching design skills is for several reasons a very complex activity. First, in order to design a product students need to understand concepts and procedures from several domains. When taught in isolation this knowledge could be compartmentalized instead of integrated in the design activity. Second, the design activity itself is usually thought to be a valuable teaching tool – 'learning by doing' – in that the students experience not only the problem and the information needed but also the strategy to solve the problem. However, because of the complexity of design problems and the heuristic character of the process involved in solving ill structured problems, it is hardly possible to give students clear and detailed methods, like algorithms, to work through the process in order to get a good design result. Descriptions of these methods in the domain of industrial design are all equally global in that it is hard to deduce from them rules that could guide design activities. This is equally true for the attempts to describe design processes in detail. For that reason they are not very useful in instruction either. So, in design

education any information about what kind of knowledge and skills are to be mastered to improve the design activity is mainly available on a intuitive level. Only recently attention is paid to the question what the understanding of knowledge in design can add to design theory and also to design education. Research on this topic has been conducted mainly in domains in which the problems are well structured, such as in physics and mathematics. Studies in these domains show that detailed domain–specific knowledge within a domain is necessary to successfully solve well structured problems inthat domain, problems that have a known solution and a known way leading to that solution, at least for experts in the domain. Central is also the idea that a suitable organized cognitive structure plays a crucial role in the quality of the problem solving process, especially in encoding and retrieving knowledge that is relevant to the problem at hand.

Empirical studies

Two studies will be presented in which we tried to measure the domain–specific knowledge design students apply during the design process, in relation to the quality of the result of the design process (a design). In the first study the 'learner reports' of first–year students were used to analyze their knowledge base. In the second study protocol analysis was used to observe the design process of students of both novice and intermediate level. In this second study the knowledge base of the subjects was related to the quality of their designs. Both methodological issues and the results of the content analyses will be discussed.

References

Alexander, P.A., Schallert, D.L., & Hare, V.C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. Review of Educational Research, 61, 315–343.

Anderson, J.R. (1987). Skill acquisition: compilation of weak–method problem solutions. Psychological Review, 94, 192–210.

Elio, R., & Scharf, P.B. (1990). Modelling novice–to–expert shifts in problem solving strategy and knowledge organization. Cognitive Science, 14, 579–639.

THE CAUSAL EFFECTS OF DOMAIN-SPECIFIC PRIOR KNOWLEDGE ON STUDY SKILLS AND CURRICULUM OUTCOMES IN PSYCHOLOGY AFTER FIVE ACADEMIC YEARS

A. Minnaert, & P.J. Janssen, Centre for Schoolpsychology, University of Leuven, Leuven, Belgium

Empirical and theoretical background

Entrance selection of freshmen isn't customary in Belgium. Due to this fact, failure rates of freshmen are relatively high (De Neve, 1991): 60 per cent in Human Sciences, 51 per cent in Biomedical Sciences en 49 per cent in Exact Sciences.

Apparently these freshmen didn't appropriately assess their potentialities in starting their career at the university level. In order to improve the psychological and educational insights into the so called "threshold" of the first year in higher education, a structural model of studying for individual differences in academic achievement is suggested.
To imitiate the study situation of a first year student, we designed a context valid study skill test which was presented as "an excursion into the domain of my future study". The underlying theoretical framework was derived from Janssen's (1989) theory of studying as the integration of learning and thinking on the base of motivation. We assume that academic performance in higher education is influenced directly and/or indirectly by a combination of cognitive and motivational variables.

Method
In the first week of the academic year 1986-1987, 169 freshmen in the Department of Psychology at the University of Leuven participated in our three hours study skill test. In september 1991, 40 % of them graduated in time, 39 % dropped out before and 21 % showed at that moment a retardation of 1 to 3 academic years. The structural model was tested on a sample of 161 students (44 men and 117 women) who participated in all the final exams at the end of their first year at the university. Our excursion involves three distinct tests. A hierarchical *silent reading test* - 10 content valid texts on topics of the study program Psychology and appropriate multiple choice questions - was designed to measure the thinking processes in deep-level learning: serialist (analyse) and holist thinking (synthesise). The students were also asked to evaluate their domain-related interest for each text separately. Thereafter a *general information test* was presented to measure domain-specific prior knowledge in psychology incidentally picked up during a well passed off process of vocational choice. Afterwards all students took an *unannounced closed* book "exam" about the silent reading texts. This "exam" intended to measure the domain-related expertise as developed while studying intentionally the texts. This latter implies a process of goal-oriented restructuring of the acquired new information into the cognitive structure already built up.
We constructured fairly reliable measures within each test (Minnaert and Janssen, 1992). Also students' curriculum completed in High School, their percentage of points obtained at the end of their first year in Psychology and their study output after five academic years was registered. To evaluate the measurement models and the structural equation model by the parameter estimation method of maximum likelihood, the computer program LISREL was used (Jöreskog and Sörbom, 1989).

Results and discussion
The hypothesed LISREL model fits well the observed data and cannot be rejected ($Chi^2 = 51.60$, $df = 54$, $p = .568$): GFI = .954, AGFI = .922, RMSR = .048.
The LISREL model explains ($R = .74$) 55 per cent of the variance in study success and study progress after five years. The structural equation parameters confirm the hypothesized effects of domain-specific prior knowledge ($p<.05$), domain-related interest ($p<.01$), curriculum completed in High School ($p<.01$) and goal-oriented restructuring study skills ($p<.01$) on study success and progress in Psychology.

Furthermore, the results confirm the significant direct effect of domain-specific prior knowledge on the speed and accuracy of study behaviour ($p<.01$) and on goaloriented restructuring ($p<.05$). So the hypothesis that the amount of present knowledge (schemata for understanding) is an important determinant of goaloriented restructuring in studying new information cannot be rejected. Glaser (1984) already pointed out that the amount of initial problem solving representations influences the efficiency and accuracy of thinking and learning within that specific domain. The importance of both declarative and procedural knowledge is already stressed by other researchers (Decruyenaere and Janssen, 1989; Weinert, 1987). Although the direct effect of domain-specific prior knowledge on goal-oriented restructuring isn't very substantial, the total effect (the sum of direct and indirect effects) is significant at the 1% level. The same conclusion can be drawn about the effect of domain-specific prior knowledge on the study success and progress after five years. Domain-specific prior knowledge as well as domain-related study skills are necessary prerequisites to accomplish study success and progress in Psychology. Besides, differences in curriculum outcomes are also significantly ($p<.01$) influenced by differences in domain-related interest.

In order to verify the substantial role of domain-specific prior knowledge and domain-related skills in studying, we examined the probability of study progress during five academic years. Therefore we used logistic regression analysis in a group of low, medium and high levels of domain-specific prior knowledge. In each logistic regression model the (lack of) study progress functions as categorical dependent variable and the structural equation variables are the independent measures. The goodness of fit of these models are quite satisfying (varying from .454 to .869) and the model improvements are all significant at the 1% level. Within the group of low level prior knowledge (75% correctly classified) differences in domain-related interest and in curriculum completed in High School turn out to be substantial predictors of study progress. Within the group of medium level prior knowledge (71% correctly classified) differences in the speed and accuracy of information processing in study behaviour are very significant. Within the group of high level prior knowledge (87% correctly classified) differences in curriculum completed in High School and in goal-oriented restructuring study skills significantly predict study progress.

Conclusion

The results of this research on studying in Higher Education, defined as the integration of learning and thinking on the base of motivation, force us into further investigations on the acquisition of structures of domain-specific knowledge in relation to domain-related study skills. On the base of its content and nomological validity, our instrument has the necessary predictive validity ($R=.74$) to promote better vocational choices of entering freshmen. It would imply lower failure rates of freshmen - given constancy of criteria - in the case the faculty should have the right to select its first year students. As yet it has not.

References

Decruyenaere, M., & Janssen, P.J. (1989). A structural model for individual differences in academic performance of freshmen. In H. Mandl, E. De Corte, N. Bennett, & H.F. Friedrich (Eds.), Learning and Instruction European Research in an International Context: Volume 2.1 (pp. 481-495). Oxford: Pergamon Press.

De Neve, H. (1991). Studierendementen aan de KU.Leuven in de Periode 1984-85 tot en met 1998-1989. (Study outputs at the K.U.Leuven in the period 1984-85 upto 1988-89.) Leuven: Dienst Universitair Onderwijs.

Glaser, R. (1984). Education and thinking. The role of knowledge. American Psychologist, 93, 93-104.

Janssen, P.J. (1989). Task, development and process in student learning: toward an integrated theory of studying. European Journal of Psychology of Education, 4, 469-488.

Jöreskog, K.G., & Sörbom, D. (1989). LISREL 7 User's Reference Guid (1 st ed.). Mooresville, Indiana: Scientific Software, Inc.

Minnaert, A., & Janssen, P.J. (1992). Success and progress in higher education: a structural model of studying. British Journal of Educational Psychology, 62, 184-192. (in press)

Weinert, F.E. (1987). Developmental processes and instruction. In E. De Corte, H. Lodewijks, R. Parmentier, & P. Span (Eds.), Learning and Instruction. European Research in an international context: Volume 1 (pp. 1-18). Oxford/Leuven: Pergamon Press/Leuven University Press.

APPLICATION OF SUBJECT KNOWLEDGE IN THE PRIMARY SCHOOL

Charles Desforges, & Steven Bristow, University of Exeter, School of Education, Great Britain

Background

Whilst it is widely held that 'old' knowledge plays a key role in the acquisition of 'new' knowledge the mechanisms of this role are not understood (Bereiter, 1985). Advances in understanding this process would contribute considerably to the design and management of classroom work.

The study reported here attempts to make such a contribution. It is part of a large project on the application of children's subject knowledge. The aims of the project are to identify how children apply knowledge (broadly conceived to include subject matter, general knowledge, knowledge of working practices and strategic knowledge) in learning from school work.

The study

The particular aim of the study was to model how children apply their knowledge of writing the English language in a novel and difficult task, specifically, the writing of headlines.

A sample of 48 children (24 x 8 year olds and 24 x 10 year olds) was recruited. Their knowledge of writing was established through observation and in-depth

interviews regarding their writing practices. The children were asked to critique their writing and to talk through the process of composition.
Their knowledge of headlines was established using interviews involving the examination of text and associated headlines prepared for the study. The children were then asked to write and critique their own headlines under two conditions:

- as a normal classroom task using VTR as stimulus;
- as an interview task using text as stimulus.

Data comprise children's protocols and the products of their work.

Analysis

Case studies are reported of models of children's normal writing and of their headline writing. The models are compared to expose processes of knowledge application and their relevance to understanding learning is examined.

Reference

Bereiter, C. (1985). Toward a solution of the learning paradox. Review of Educational Research, 55, 2, 201–226.

TEACHING INDUCTIVE THINKING SKILLS TO RETARDED AND NORMAL STUDENTS

F.P. Büchel, G. Pelgrims, U. Scharnhorst, University of Geneva; P. Büchel, University of Fribourg Switzerland, & J. Hamers, University of Utrecht, Utrecht, The Netherlands

The scientific analysis of retarded performance focused for the last twenty years on metacognitive knowledge and executive control (Borkowski & Büchel, 1983). Although there is no doubt about the importance of meta-cognitive functions, we believe that the analysis of the actual reasoning activity, especially of inductive reasoning, has been neglected. Although metacognitive research has demonstrated that low-functioning students can acquire declarative and procedural knowledge, they do not necessarily transfer it to new situations (Campione & Brown, 1984). Understanding the cognitive deficiencies of this population increasingly becomes a question of understanding their lack of transfer ability. Transfer of a successful strategy from one situation to another can be viewed as an inductive process. Inductive reasoning is considered to be the main component of general intelligence (Snow, Kyllonen, & Marshalek, 1984). Taking these arguments into account, the study of inductive reasoning in low-functioning students seems to be promising with respect to practical and theoretical reasons. From a practical point of view, this research may have implications for cognitive education as well as for dynamic assessment. If strategy transfer is based on inductive processes, cognitive education of retarded students should focus on inductive reasoning. With respect to assessment practice, dynamic testing offering specific help to the student, seems to be more valid than static intelligence testing. This is especially true for low-functioning students (Guthke, 1990). A better understanding of the optimal help conditions could increase the reliability of dynamic testing.

From a theoretical point of view, it would be helpful to clarify the relation between training conditions and transfer effects. This is especially important with respect to retarded students. A significant relation between general ability level and transfer of inductive training could be demonstrated (Ferrara, Brown & Campione, 1986). This research also showed that the ability-transfer relation might be confounded with different help conditions. We therefore postulate an ability x help condition x transfer interaction. Help hierarchies of several dynamic tests (e.g. Ferrara et al., 1986, Carlson & Wiedl, 1978) consist of three different kinds of help: Simple feedback, restructuring of the task and verbalization of the implicit rules. These help conditions correspond to three important positions in educational theory:

behaviorism, constructivism and more recent analytic approches (e.g. Klauer, 1988).
In order to test the outlined research questions, a training experiment was conducted. Normal and retarded students, matched on mental age and pretest performance, are trained on inductive reasoning tasks. Learning and transfer effects are evaluated with an inductive reasoning test combining three task types, three complexity levels and five modalities of presentation. The test structure allows to define certain subsets of items as evaluating:

- Learning effects: these items are structurally identical to the training tasks which require the completion of concrete and abstract figural series;
- Near transfer effects: these tasks differ from the training tasks with respect to task type, or complexity level, or modality;
- Far transfer effects: these tasks differ from the training tasks in more than one dimension.

Experimental subjects were attributed to one of three different training conditions:

- Condition I: Simple feedback;
- Condition II: Restructuring of the task;
- Condition III: Restructuring of the task and explicit verbal formulation of the inductive rules.

With respect to learning and transfer effects the following hypotheses were formulated:

- Learning effects;
  Normal students will profit from all three conditions. Retarded students will not profit from condition I, but from conditions II and III.
- Near transfer effects;
  Normal students will profit from all three conditions. Retarded students will not profit from condition I, but from conditions II and III.
- Far transfer effects;
  Normal students will not profit from condition I, but from conditions II and III.
  Retarded students will not profit from conditions I and II, but from condition III.

The hypotheses are tested with a factorial design:

- Independent factors;
  Subjects (2): normal (9-10 years); retarded (12-16 years).
  Training conditions (3): systematic correct-incorrect feedback; feedback and restructuring of the task; feedback, restructuring and explicit formulation of the inductive rules
- Dependent factors;
  Test tasks (learning, near transfer, far transfer)

The 18 training tasks are presented over three training sessions (3x6 tasks). All tasks as well as two introductory exercises are presented in a multiple-choice format on a personal computer AMIGA 2000. Experimental data will be presented and discussed. Our main interest focuses on the discussion of differential transfer effects. With respect to the inductive reasoning test, data of the statistical task analysis as well as normalisation data will be presented for normal and retarded

students. We emphasize the differentiation between task types, complexity levels and modalities as well as problems of reliability and validity. For retarded students, a test-retest reliability and validity studies will be presented.

References

Borkowski, J.G., & Büchel, F.P. (1983). Learning and memory strategies in the mentally retarded. In M. Pressley & J.R. Levin (Eds.), Cognitive strategy research. Psychological Foundations (pp. 103-128). New York: Springer.

Campione, J.C., & Brown, A.L. (1984). Learning ability and transfer propensity as sources of individual differences in intelligence. In P.H. Brooks, R. Sperber, & C. McCauley (Eds.). Learning and cognition in the mentally retarded. Hillsdale, NJ.: Erlbaum.

Carlson, J.S., & Wiedl, K.H. (1978). Use of testing-the-limits procedures in the assessment of intellectual capabilities in children with learning difficulties, American Journal of Mental Deficiency, 82, 559-564.

Ferrara, R.A., Brown, A.L., & Campione, J.C. (1986). Children's learning and transfer of inductive reasoning rules: Studies of proximal development, Child Development, 57, 1087-1099.

Guthke, J. (1990). Les tests d'apprentissage comme altérnative ou complément des tests d'intelligence. Un bilan de leur évolution. In F.P. Büchel, & J.-L. Paour (Eds.), (1990). Assessments of learning and development potential: Theory and practices. European Journal of Psychology of Education, (Special issue), 2, 117-133.

Klauer, K.J. (1988). Paradigmatic teaching of inductive thinking. In E. DeCorte et al. (Eds.), Learning and instruction Oxford: Pergamon Press.

Snow, R.E., Kyllonen, P.C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R.J. Sternberg (Ed.), Advances in the psychology of human intelligence (47-103). Hillsdale, NJ: Erlbaum.

DEVELOPMENT OF A COMPUTER-ASSISTED INDUCTIVE REASONING TRAINING PROGRAM

P. Büchel, University of Fribourg, Switzerland

On a personal computer AMIGA 2000, three different implementations of multiple choice series completion tasks were developed. Every task involves two relevant attributes extending over periods of different length. The implementations differ in the level of explicity of the computer-presented hints for task solution. After a short introduction, the program can be used without any assistance. The tasks are presented one after another in a fixed order. The answer is selected with the help of a mouse.

In condition I, response selection is followed by a simple correct-incorrect feedback and, in the case of an incorrect choice, by the instruction to try again. In condition II, hints are presented on four different levels of explicity. After a first incorrect choice, the subject's attention is focused on the first relevant attribute of the task (e.g. form). After a second incorrect choice, both attributes are shown (e.g. form and colour). A third incorrect choice leads to the presentation of the periodical

structure of the first attribute. After the fourth incorrect choice, both attributes and their respective periodicity are presented. Finally, the correct answer is given. All hints consist of nonverbal focusing of task relevant components. In condition III, nonverbal hints are additionally explained by an explicit verbal formulation of the relevant inductive rules.
All instructions are presented in a written as well as in a spoken format, the hints are additionally animated with elements of computer graphic. In the more elaborate training conditions (II and III), every confirmation of a correct answer is followed by a summary of all relevant hints. A background program records all mouse-movements and mouse-clicks. This protocol can be used in its symbolic form, but it also allows the visual and auditive replay of the session.The structure of the program together with some examples of training tasks and hint sequences will be presented.

CONSTRUCTION AND EVALUATION OF AN INDUCTIVE REASONING TEST

G. Pelgrims, University of Geneva, Switzerland

A main problem in inductive reasoning research consists of the lack of appropriate dependent measures with respect to differential transfer effects. Researchers mostly refer to subscales of intelligence tests. Complexity and difficulty levels are rarely controlled, and a systematic variation of modalities of presentation is almost impossible.
In order to evaluate learning and transfer effects of an inductive training, a paper and pencil test has been constructed combining three task types (classifications, series, linear analogies) with three complexity levels (low, medium, high), and five modalities (figurative-abstract, figurative-concrete, verbal-abstract, verbal-concrete, symbolic). With respect to complexity levels, three different complexity models have been elaborated, one for each task type. The test consists of 45 items, every combination of the three dimensions being realized by one item. The test contains therefore 15 items of each task type and each complexity level, and 9 items of each modality. Task analyses were conducted with 112 normal and 82 learning disabled and mentally retarded students. Matched on mental age, normal students performed better than retarded students in all three task types. The three complexity levels were reflected in respective mean differences in both populations. With respect to modalities, the results are less clear (Pelgrims, 1991).
A test-retest-reliability study as well as first estimations of validity have been undertaken with 80 retarded students. Standardisation data have been collected with 500 normal and 400 retarded students.
Principles of test construction as well as the already available data will be presented and discussed.

References

Pelgrims, G. (1991). Construction d'un test de pensée inductive: de l'élaboration des items à l'analyse de la difficulté et de la sélectivité. Travail de mémoire soutenu à la Section des Sciences de l'Education de l'Université de Genève.

HOW TO TRAIN INDUCTIVE REASONING SKILLS? THE INFLUENCE OF THE LEVEL OF EXPLICITY OF TRAINING ON LEARNING AND TRANSFER EFFECTS IN NORMAL AND RETARDED STUDENTS

U. Scharnhorst, University of Geneva, Switzerland

Inductive reasoning has proved to be the main factor of general intelligence. In order to understand the cognitive functioning of retarded persons, the study of inductive reasoning skills should therefore be a promising approach. The fact that students with low intelligence perform poorly on inductive tasks is well known and represents a somewhat circular argument. It is more interesting to clarify the question of the source of this impairment. It can be argued that retarded students are inefficient in inductive reasoning because their educational environment did not foster these skills. If they are explicitly trained on well structured inductive tasks, they should be able to overcome this deficiency.

This hypothesis was tested in a training experiment. Normal and retarded subjects were attributed to one of three different training conditions involving three individual training sessions with concrete and abstract figural series completion tasks. The tasks are presented in a multiple choice format on a personal computer. In training condition I, the subjects work independently on the tasks and simply receive a systematic correct-incorrect feedback after each answer choice. In condition II, the selection of an incorrect answer is followed by error-specific visual hints. Condition III corresponds to condition II, but it is supplemented by an explicit verbal formulation of the relevant inductive rules. Five normal and five retarded subjects, matched on mental age and pretest performance, were attributed to each training condition. Two control groups only participated in the pre- and posttest sessions. Pre- and posttests consisted of an inductive reasoning test which was administered before, directly after and a month after the training sessions. Because the test combines three task types with three complexity levels and five modalities of presentation, it is possible to observe transfer effects from training in a differenciated way.

Experimental data will be presented and discussed. Our main interest focuses on the discussion of differential transfer effects. As we have also collected reaction time data with a computer background program, a systematic analysis of learning effects within and between sessions should be possible.

RECENT DEVELOPMENTS IN INSTRUCTIONAL DESIGN

Prof.dr. P.R.J. Simons, Catholic University of Nijmegen, Nijmegen, & J. van Merriënboer, Department of Education, University of Twente, Enschede, The Netherlands

Summary

Traditional instructional design theories and models are developed from behavioral theoriesand from general models how to solve an educational problem. Based on ideas and researchresults of cognitive science and the development of highly interactive, technology baseddelivery systems it became clear that the traditional design theories and models had somelimitations. The most important criticism is the analytical nature of instructional designtheories. All theories start with identifying the components and subcomponents of thesubjectmatter which then are used for designing and sequencing instructional frames. Especially incase of complex subject matter which contains several concepts, principles and proceduresthestudent may be unable to integrate the components in organised wholes. The analyticalnatureof the traditional theories also causes an economic problem, because the designer ofinstruction needs to compose every instruction from the basic elements, which is laborintensive. Finally traditional instructional design theories mainly refer to tutorial instructionin which the student is supposed to be passive.

The symposium addresses the limitations and criticisms and will discuss how recentlydeveloped instructional design theories and models which are embedded in or derived fromcognitive science try to solve the instructional design problems. Attention is given to inquiryteaching, to the constructivist approach, to the use of contextual knowledge and to theintegration of knowledge and skills through problem solving techniques. The application ofcognitive orientations in designing exams also will be addressed.

Introduction

Traditional instructional design models are developed from behavioral theories and fromgeneral models how to solve an educational problem. Based on the ideas and findings ofcognitive psychology in the seventies and early eighty's new models and design rulesemerged, which roughly comprise four approaches: inquiry teaching (Collins, 1977,1983),instructional design strategies for concept teaching (Tennyson and Cocchiarella, 1986),constructivism (Brown, Collins and Duguid, 1989; Jonassen, (1992)) and the use ofcontextualknowledge (Brown, Collins and Duguid, 1989; Jonassen, (1990) & Tennyson (1992).

NP 425

The purpose of the inquiry method is teaching students to reason from cases in order todevelop a knowledge network. The results of experiments on human memory and reasoningsupported the supposed structure of human knowledge as a rich semantic network whichcomprises declarative, procedural and metaknowledge. The knowledge base consists ofincomplete and inconsistent knowledge. The students use this knowledge in reasoning and indrawing conclusions. Collins, Warnock, Aiello and Miller (1975) analyzed the reasoning ofstudents, especially the negative and functional inferences and published the first results ofa method of inquiry teaching. Collins (1977, 1983) further developed the rules of inquiryteaching and presented a theory of inquiry teaching. The theory had three parts: the goals ofteachers; the strategies teachers use and the control structure that governs their teaching. Themethod of inquiry teaching was described independent of the content of the subject matter,but several publications make clear that subject matter which shows causal structures isoftenused to illustrate the method of inquiry teaching.

Tennyson and Cocchiarella (1986) presented an empirically based instructional design theoryfor teaching concepts. The theory was un update of the Merrill and Tennyson (1977) model,which was based on the classical theory of concept learning. Evidence for this theory wasmainly derived from carefully designed laboratory experiments with artificially constructedstimuli which could be categorized with certainty once the concept was learned. In theseventies the classical theory faced several problems such as the existence of disjunctiveconcepts, the existence of unclear cases and the failure to specify defining features for mostconcepts. In research attention was given to the learning of natural concepts and conceptswithless "well-defined" attributes. The results indicated that the process of concept learning hastwo phases: the formation of conceptual knowledge (prototype formation) and thedevelopment of procedural knowledge. Conceptual knowledge is formed by the integratedstorage of meaningful dimensions (defining and variable attributes) selected from knownexampies and by the connection of this information to a given domain of knowledge. Theinstructional design model which was developed by Merrill and Tennyson emphasized theinstruction of the defining attributes and the instruction of an isolated concept. TennysonandCocchiarella however elaborate on prototype formation by defining a "best example" in theirinstructional design strategy. They also pay attention to attribute characteristics (constantandvariable) and to the relational structure of concepts (successive and coordinate). Thesefactorslead to attribute elaboration and to the successive and simultaneous presentation of examplesas components of the instructional strategy. The importance of prerequisite knowledge isunderlined for which embedded refreshment is a necessary instructional design component.

Current cognitive theory emphasizes that learning is a process of knowledge construction,notknowledge absorption. Learning occurs by interpreting information and by checking whetherthe supposed consequences will happen. Effective learning depends on the intentions, self-monitoring and representational constructions of the learners. Thus the instructiveenvironments should be designed in such a way that they provide information and questionsfor the knowledge construction processes of the students. They should support multipleperspectives of reality, knowledge construction and context-rich

experience based activities(Jonassen, 1991). The important application of the constructivistic view of learning is theprovision of instruction in relevant contexts. Learning should occur most effectively incontext. The context becomes an important part of the knowledge base associated with thatlearning. Brown, Collins and Duguid (1989) stressed the distinction between authentic andschool activities. When authentic activities are transferred to the classroom, the context ischanged and the tasks become classroom tasks instead of real tasks. As a result, they state,conceptual and problem-solving knowledge acquired in school remains largely unintegratedor inert for many students. Learning should be adjusted to the situation in which it takesplace. Instruction should provide for a relevant context. The contributions of the symposium elaborate on the changes in instructional design theoriesand how the changes influenced the actual design of instruction. The results of the researchaddresses the effects of the instructive environments, both retention and transfer.

THE INSTRUCTIONAL DESIGN, STUDENT ACTIVITY, DISCOVERY LEARNING AND TEACHING TO SOLVE PROBLEMS

S. Dijkstra, Department of Education, University of Twente, Enschede, The Netherlands

Summary

Recently traditional instructional-design theories and models, which are developed frombehavioral learning theories and from general educational problem-solving models, arecriticized. The traditional theories and models were too analytic and unable to teachintegratedwholes. The analysis of the subject matter or the target behavior of the students usuallyresulted in different outcomes of instruction such as facts, concepts and principles. For eachof these categories a further analysis was made and the components of the outcomes wereformulated such as attributes and relations. The results of the analysis were used to designand sequence the instructive frames. The emphasis was on the presentation of information orexternal conditions for learning rather than on what the student should do. Thereforetraditional or first generation instructional design was only seen as a tutorial approach whichsupposed the students to be passive and which was insufficient to design the teaching ofproblem solving activities. This paper will try to link the knowledge of human problem solving with recentdevelopmentsin instructional design. It will address the issue of the analysis of the target objective, howdetailed this should be done and how the resulting components can be used to formulatequestions and problems which stimulate the student's activity, enhance the formation ofcognitive networks or schemata and in some situations lead to discovery of problem-solvingprocedures. The tutorial approach will be compared with an experiential approach that fordifferent types of content can realise continuous transactions with subject matter. Attentionwill be paid to understanding and applying knowledge in solving complex problems.

REFLECTIONS ON APPLIED INSTRUCTIONAL DESIGN IN TEST CONSTRUCTION AT THE MACRO AND THEMICROLEVEL: DESIGN AND IMPLEMENTATION OF EXAMINATION "VERSICHERUNGSFACHMANN/-FACHFRAU (BWV)"

K. Breuer, University of Paderborn, Gütersloh, Germany

Summary

Educational design in general aims at the layout of learning environments. This usually doesnot include test construction and development, though evaluation of learning outcomes is astandard part of learning-teaching activities.

While it is obvious that the orientation towards cognitive learning-teaching theories hasimplications for the design of learning environments there is not much considerations yet inrespect to the design of examinations.

The introduction of the new basic qualification "Versicherungsfachmann/-fachfrau (BWV)"/ insurance expert / for insurance agents in Germany has given the opportunity to refer tocognitive orientations in educational design within the development of the correspondingexamination. Cognitive-based principles have been applied to the lay-out of the system onthemacro as well as the micro level.

The design principles are based within the distinction ofdifferent kinds of learning outcomes, i.e. declarative, procedural, and contextual knowledge,and test for these.

At the micro level specific procedures are used for writing the test items, and for the conduction of the "advice and sale" part.

Beyond the report on the design principles the contribution will report on the experiencesfrom the implementation of the system at the national level within 15 examination centersthroughout Germany, with more than 600 newly trained examinators and several thousandexaminees at the time of the conference.

COGNITIVE PSYCHOLOGY APPLIED TO THE DESIGN OF PRINTED STUDY MATERIALS AT THE UNIVERSITY LEVEL

J. Lowyck, Catholic University of Leuven, Leuven, Belgium

Summary

Instructional design (ID) as a field of research recently has been confronted with challengesfrom cognitive psychological theories. After the information-processing approach, theconstructivistic paradigm questions traditional ID-models which primarily focused onpresentation variables. The application of cognitive psychological findings to the design ofprinted study materials implies the need for an emphasis on the conceptual structurereflectedin these materials on the one hand and the (hypothetical) mental models or cognitivestructures of students on the other, in order to enable students to activate all necessarylearning processes. Moreover, ID-models need empirical validation as to their cognitivequality.

In this paper it is assumed that the validation of theoretically sound ID-principles impliestheirapplication in ecologically valid settings. In the study reported here, the effectiveness of twoversions of a self-study package on "General Didactics" are compared in order to investigatewhether interventions from a cognitive perspective toward learning lead to enhancedlearning.The first version was written by domain-experts and aimed at clarifying as much as possiblethe concepts and their interrelationships. Schema's of the conceptual structure of all parts ofthis self-study package were made by the experimenter. These schema's were, then,analyzedin order to detect possible students' problems with concepts and relations between concepts,that would hamper the application of effective learning strategies. This analysis completedbyproblem indications and comments by students, was the starting point for revising the self-study package. This revision resulted in the second version of the self-study package.
Problems and comments of students were (overall) in accordance with the problems asindicated through the analysis by the experimenter on the basis of cognitive learningprinciples. The revised version, however, did not necessarily result in an improvement of alltypes of learning results, as indicated by the comparison of examination results of studentsstudying one of both versions of the self-study package.

DIFFERENT SUBJECTS, DIFFERENT CONTEXTS , DIFFERENT LEARNERS, - ONE ID-MODEL? A RECENT AND OLD CHALLENGE TO INSTRUCTIONAL THEORY

F. Schott, Deutsches Institut für Fernstudien, Universität Tübingen, Tübingen, Germany

Summary
Designing instruction is an ambitious venture. The problem requires a decision how to teacha certain subject in a certain context certain learners. Taking into account that there is a vastamount of different subjects, of different contexts, and of different learning individuals, themultiformity scale of this problem's conditions runs into infinity. From this point of viewtheproblem of designing instruction seems to be more an ambiguous adventure than anambitiousventure, especially if the intention is to solve the problem with one single instructionaldesign(ID) model. Otherwise it would not be helpful to have as much ID-models as the amount ofmultiformity of the different ID-problems' conditions. Behavioristic approaches offeredassistance to reduce the vast amount of varieties by using conceptually simple theoriesdrivenby the corresponding ideal of physical theory building. Whereas modern cognitivepsychologypoints out successfully that the learning process has to take into consideration differentstructures of subject matter, different contexts and different learners. Recent conceptscharacterize this trend, like e.g. "contextual knowledge", "situated learning", "metacognitionand self-directed learning", "constraints of knowledge and cognitive development", "differentknowledge-structures of experts and novices" and last but not least "constructivism in ID".Therefore, compared with the times of behaviorism the

question "Different subjects,differentcontexts, different learners, - one ID-model?" has aggravated to a recent challenge toinstructional theory. In my contribution I try to discuss this question against the backgroundof new developments of ID-models. Are there concepts which can help to attack theproblemin question? Are there new answers to this old challenge to instructional theory?

PREDISPOSITION OF LEARNERS AND ITS INFLUENCE ON LEARNING

P.A. Kirschner, F. Dochy, P.R.J. Simons, The Netherlands, & E. De Corte, Belgium

THE INFLUENCE OF PRIOR KNOWLEDGE ON LEARNING: A QUESTION OF ECOLOGICAL VALIDITY OF RESEARCH AND DIFFERENT PRIOR KNOWLEDGE COMPONENTS?

Filip J.R.C. Dochy, Centre for Educational Technology and Innovation, Heerlen, The Netherlands

Introduction

The role of prior knowledge in learning is generally recognized as being important. The generally recognized experimental research results as shown by studies of Bloom, Weeda, Lodewijks, Haertel, Walberg, et. al (for an overview, see Dochy, 1992), present a set of convincing figures ranging from 60 to 95 % of variance in post-test scores explained by prior knowledge. Some of these studies however lack ecological validity. The questions raises why prior knowledge state tests are not yet implemented in school learning as a main determinant of the learning process if the above are transferable to real life learning.

The paper reports two investigations using different prior knowledge state tests, being domain-specific, subject-oriented of cross-domain tests. The results show that different kinds of prior knowledge explain different amounts of variance in post-test scores and relate to different components of a students' prior knowledge. The explained variance ranges between 16 and 42 %.

Prior knowledge research

The fact that prior knowledge turns out to be an important educational variable to account for post-test variance was shown in several investigations (for an overview, see Bloom (1976) and Dochy (1992)). Weeda (1982) found that knowledge measured prior to a course, explained, on average, no less than 50% of the variance in the post-test scores. Comparable results were reported by Bloom (1976) who found correlations between 0.50 and 0.90 between pre-test and post-test scores. From these correlations, Bloom deduced the amount of explained variance. Lodewijks (1981) found a correlation of 0.60 between the assessment the students made of their prior knowledge and their performance in post-tests. The results of the research into the influence of variables on study results demonstrate that prior knowledge explains between 30 and 60 per cent of the

variance in study results. Studies aiming at causal modelling, using LISREL or LVPLS techniques, and reaching a good overall fit and a multitude of significant structural coefficients, stress once more the importance of prior knowledge (Parkerson, et.al, 1984; Körkel, 1987).

Ecological validity

However, some remarks concerning the research cited in the section above can be made.

It is striking, that in the research referred to, activation of prior knowledge is often used but little use is made of the prior knowledge itself. In experimental situations, a short text or task is offered as an independent variable. Perhaps it is more interesting to look at the impact of students' real and present prior knowledge on the learning process and learning results. Further, research into prior knowledge is often characterized by a limited ecological validity. This means that the experimental environment is so constructed that the research results cannot be applied immediately to real educational situations. Finally the nature of the test (general questions or specific questions on the text, recognition or remembering, etc.) is seldom taken into account in the interpretation of the results.

Domain-specific prior knowledge

Recent studies, trying to cope with the disadvantages of a limited ecological validity, have tried to scrutinize the effect of prior knowledge in real educational settings. These studies, such as the Minnaert & Janssen study (1990) and the Weinert studies (1989) have been investigating the domain-specific prior knowledge.

Recent research has pointed to the fact that both domain-specific knowledge and domain-transcending knowledge exist in the knowledge base. Furthermore, there is evidence that learning is far more domain-specific than earlier theorists of learning believed, i.e. concrete and practical situations seem to be better learning environments than highly abstract ones. Carey (1985) suggests that the acquisition of knowledge during the total period of development i.e. throughout a person's life, is based on increasing knowledge within various domains. This 'domain-specific restructuring view of development' has received a great deal of support in research on novice-expert differences within various domains.

There is a considerable amount of evidence that the domain-specific prior knowledge is the form of prior knowledge that most affects the learning process and results. The importance of domain-specific knowledge has been well demonstrated for students of different ages (Glaser, 1987; De Corte, 1990b).

Further, the results from studies on metacognition show remarkable parallels with the results from intelligence studies looking at predictors of learning outcomes (Weinert, 1989). Schneider (1985) found a significant correlations, but Weinert (1989) replicated the studies, partialling out the effect of prior knowledge, and concluded that the former conclusions did not hold any longer. The correlations between metacognition and performance diminished just as much as in the intelligence case i.e. a decrease between .34 and .21. (r was between .07 and .20). Even on the basis of other studies Weinert (1989) concluded that "contrary to expectations, past research has shown that motivational variables and instruc-

tional characteristics make very little contribution to the prediction of school performance".

Investigations using Prior Knowledge State Tests: the influence of knowledge state components

The research described is an attempt to grasp a student 's domain-specific prior knowledge state and its impact on learning, by concentrating initially on the construction of a set of prior knowledge state tests and paying attention to the role of different prior knowledge state components. In past research, existing course-related tests were mostly used to assess prior knowledge (De Corte, 1991), without differentiating between types of the prior knowledge state along certain dimensions (Dochy, 1990).

This investigation focuses on the subject-oriented prior knowledge state (SO PKS) and cross-domain prior knowledge, i.e. the optimal requisite (OR) prior knowledge state, and the mathematics (MA) prior knowledge state.

Based on our earlier investigations, we expect that higher scores on subject-oriented and cross-domain prior knowledge state tests will result in higher scores on a post-test. We do not expect differences between student types (ES and LS), following the results of ex post facto research project 2 (Dochy, Bouwens, et. al, 1991). Nevertheless, we repeated the analysis with the 'student type' variable, even though the current research focuses primarily on the quality of the prior knowledge state and its impact in terms of relationship to post-test scores.

Research procedure and data analysis

The research procedure consisted of several phases: Registration and introductory session; administration of the four prior knowledge state tests [SO KST1 (± 20 min.),OR KST (± 45 min.),ME KST (± 15 min.),MA KST (± 45 min.)]; a first study period (± 45 min.): during this study period the students studied the text of learning units 14 & 15 of the "Economics & Money" course ; lunch time (30 min.); a second study period (± 45 min.); administration of the SO KST2 post-test (± 15 min.); concluding session. The data were analyzed using analysis of variance and regression analysis techniques.

Conclusions

The first conclusion is that the variable "student type" (ES or LS) has not been helpful to detect differences in overall PKS within a population of students. However, these
differences could be extrapolated in terms of specific components of the prior knowledge state. The differences in mathematics prior knowledge state and optimal requisite prior knowledge state between both sub-populations were striking. These facts call into question the multi-functional nature of the "Economics and Money" course and imply a need for structural, organisational or educational adaptations of this course.

The second conclusion of this research is that it could be stated that the level of PKS is helpful to predict future learning results of students. The second general measure of the PKS (PKST2), comprising the optimal requisite test scores and the test scores for the subject-oriented PKS test, explains up to 42% of the variance in the post-test scores. Further analysis revealed that the optimal requisite prior

knowledge (highest predictive power) and the mathematics prior knowledge are important components of the student PKS.
Further, it was shown that the estimation of PKS level through self-assessment by the students is not very reliable. The estimations do not reflect the levels of PKS as measured by the objective tests. For our purpose, it can be concluded that self-assessment is not useful, at least not for introductory courses, as also shown by Falchikov and Boud (1989).

References

Dochy, F.J.R.C., Bouwens, M.R.J., Wagemans, L.J.J.M., & Niestadt, D.W. (1991). The role of subject-oriented expertise. A study of the impact of personal and contextual variables on success in an economics course as indicators of expertise: ex post facto research 2. Otic Research reports 25. Heerlen : OTIC.

Dochy, F.J.R.C. (1990). Mapping prior knowledge or expertise. OTIC Research Reports 28. Heerlen, Educational Technology Innovation Centre.

Dochy, F.J.R.C., & Bouwens, M.R.J. (1990). The construction of knowledge state tests, knowledge profiles and the measurement of the value added. OTIC Research Reports 27. Heerlen, Educational Technology Innovation Centre.

Dochy, F.J.R.C. (1990).Instructional implications of recent research and empirically based theories on the effect of prior knowledge on learning. In: Pieters, J.M., Breuer, K.,Simons, P.R.J.(Eds). Learning Environments: Recent research in Psychology. Berlin New York, Springer Verlag; pp.339 - 357.

Dochy, F.J.R.C., Valcke, M.M.A., & Wagemans, L.J.J.M. (1991). Analysis of the Quality and Impact of Expertise in Economics. OTIC Research Reports 26. Heerlen: Centre for Educational Technology and Innovation.

Dochy, F.J.R.C. (1992). Assessment of prior knowledge as a determinant for future learning. Utrecht/London: Lemma B.V./Jessica Kingsley Publ. (With preface by Prof. Robert Glaser and Prof. Erik De Corte).

Dochy, F., Valcke, M., & Wagemans, L. (1991). Learning economics in higher education: an investigation concerning the quality and impact of expertise. Higher Education in Europe, 4, 123-136.

Dochy, F.J.R.C., & Valcke, M.M.A. (1992). Knowledge profiles of economics and law students: an in-depth analysis of the prior knowledge state. OTIC research report 34. Heerlen: Centre for Educational Technology and Innovation.

Valcke, M.M.A., & Dochy, F.J.R.C. (1992). Knowledge profiles of students with low and high prior knowledge states. OTIC-research report 35. Heerlen : Centre for Educational Technology and Innovation.

Wagemans, L.J.J.M., & Dochy, F.J.R.C. (1991). Principles in the use of experiential learning as a source of prior knowledge. Distance Education, 12, 1, 86-108.

AGREEMENT BETWEEN STUDENT EXPECTATIONS, EXPERIENCES AND ACTUAL OBJECTIVES OF SCIENCE PRACTICALS

P.A. Kirschner, M.A.M. Meester, E. Middelbeek, & H. Hermans, Open University, Heerlen, The Netherlands

The Open university of the Netherlands (OuN), an institution for open higher distance education, minimizes traditional laboratories in its courses by using interactive videodiscs (IV), computer simulations and home experiment kits. A complication in using these alternatives is that for them to be successful, a student must encounter the objectives that developers put into them. Whether this occurs is at least partially affected by a student's expectations before entering the practical situation.

Theory

Students possess well-established beliefs/preconceptions when entering instruction (Confrey, 1990) which determine what they perceive and how they perceive it. But if these preconceptions are incorrect, they become *misconceptions*. Recent research on scientific misconceptions (Berg & Brouwer, 1991; Gilbert & Watts, 1983; Shuell, 1987; Hodson, 1990) has shown that students have conceptions of the natural world which are different from those of scientists. They, in turn, interpret observations in terms of their misconceptions. If this is ignored, it leads to a type of *wrong* or *alternative science* in which new data are (technically) correctly and logically interpreted, but based upon flawed premisses. For those who develop educational materials or provide educational experiences, it is imperative not only to understand what students know, but also to determine what they believe.

Suchman (1966) posited that the meaningfulness of a (sensory) encounter depends on the relevant information the learner possesses. Existing knowledge is, thus, critically important for learning. Knowledge consists of an integrated structure of conceptions plus procedures for using that structure. Conceptions are used to make sense of new phenomena, construct other more appropriate conceptions and solve problems (Hewson, 1980). In education, ideas and concepts are brought to bear on experiences, not derived from them. What the learner expects to encounter determines what she/he sees. Theobald (1968) stated: "Experience does not give concepts meaning, if anything concepts give experience meaning". Paraphrased: Experiences do not give rise to expectation, if anything expectations determine experience.

This is similar to the faulty idea of an *inductivist habit of thought* (D'Amour, 1979) which presumes that science begins with a solid base of brute facts, prior to and independent of theories. There is, however, no such thing as a 'brute' fact; all observation is theory dependent. Scientists observe phenomena selectively and the tool of this selection is the theory. Just as theory determines what scientists see and how they describe it, expectations determine what students experience and how they experience it.

Cognitive dissonance theory may also help explain why those anticipating things encounter them while those not anticipating them do not (Festinger, 1957). A

dissonant situation is unpleasant Those experiencing this strive to rid themselves of or reduce that feeling. If a student does not expect to encounter something in instruction, but comes across it anyway, she/he will try to negate this discrepancy (and accompanying unpleasant feelings) by resisting the notion that it was actually present. Thus, what *is not* anticipated *is not* encountered while what *is* anticipated *is* encountered.

The following reports on an investigation of how OuN students experience different types (laboratory and IV) science practicals. Central is: (1) the objectives experts say are explicitly covered in the practicals (actual content), (2) the objectives students feel they encountered in the practicals, (3) how the actual contents compare to the experiences of students, (4) the objectives students expect to encounter in the practicals, and (5) a comparison of expectations and experiences. The hypothesis of interest here is: What a student does not anticipate encountering will also not be encountered by her/him.

Method

Experts, those who developed a course or tutor it, evaluated four OuN practicals for the explicit presence of 64 practicals' objectives (Kirschner & Meester, 1988). Students were asked, prior to the practicals, to note which of these same learning objectives they anticipated encountering. They were also questioned subsequent to completion of the practicals which objectives they encountered. The practicals investigated were from: (1) *Form and function in biology* (Biology I), (2) *Chemistry - Theory and practice*, (3) *Physics*, and (4) *Biology of populations and behaviour* (Biology II)[1] .

Results

Table 1 presents an overview of the main results. Of note is the similarity between objectives anticipated, encountered and judged by experts as being present.

Table 1

Frequencies of present versus encountered objectives and anticipated versus encountered objectives for the four practicals studied

Practical	present versus encountered objectives		anticipated versus encountered objectives	
	common[2]	noncommon	common	noncommon
Biology I	28	11	31	5
Chemistry	25	10	32	2
Physics	22	12	30	5
Biologie II	15	16	19	7

1 The first three had traditional four day laboratories, the last one four had interactive videodiscs.

2 A common objective is one which is either considered by both students and experts as being present or is both encountered by students. A noncommon objective is one which is either considered as being present by the students but not by the experts, is anticipated but not encountered by students or vice versa.

In comparing objectives in the laboratory practicals in *Biology I, Chemistry* and *Physics* with those in the IV-practical in *Biology II*, seven objectives ascribed to the IV practical are not present in any lab practical. All of those objectives are concerned with higher order academic skills. However, six of them were neither anticipated nor encountered by students. The *Biology II* (IV) practical also distinguishes itself in that there are six objectives ascribed to the lab practicals which are not present in the IV-practical, all rather low level skills necessary during experimentation.

Because most students will have had less experience with IV-practicals than labs, it is interesting to compare their expectations to what is actually present. For *Biology I, Chemistry*, and *Physics* there are respectively 8, 7, and 7 objectives not anticipated practicals but considered present. This number jumps to 11 for *Biology II*, 8 of which are required prior to experimentation. It is fairly apparent that students are less sure what to expect from an IV-practical than from a 'traditional' laboratory.

To test the hypothesis that what a student doesn't anticipate encountering will not be encountered X^2-scores for the separate practicals and for all four of the practicals together were calculated. The scores for all four of the practicals apart and the four practicals as a whole all proved to be significant. It is clear from these results that what is anticipated and what is encountered are very closely related.

The data was also analyzed where there was a discrepancy between what experts say is present and what students say they encountered in the practicals. Again, what is anticipated and what is encountered were highly related, regardless of the nature of the discrepancy between experts and students (X^2=28.81, p<.001).

Discussion

The most important conclusion is that students' expectations unequivocally influence what they encounter. This influence is independent of what actually is present.

This result has ramifications for both educational research and educational design. For research, these results may hold the key to the apparent inconclusiveness of comparative media studies. Our results mean that any (non)effect found in such research can be as easily ascribed to students' expectations as to the media or educational techniques used. First, a student encountering new media/techniques may have misconceptions as to their possibilities and thus not be optimally aware of what they are capable of achieving. In extreme cases this can lead to a student expecting something totally different than what she/he gets. Second, a student may interpret new techniques (e.g. group work with the teacher as participant) to be old ones (e.g. group work with the teacher as expert). These misconceptions influence what a student anticipates and thus what she/he encounters. This was evident for the IV for *Biology II* where the *common quality* was much lower than the *common quality* of each of the other lab practicals. Assuming that if a student does not encounter an objective then she/he will also not achieve it, then it is not unimaginable that she/he will do worse on a criterium measure than a student using a medium or technique with which she/he was well acquainted.

The effect of this research on educational design and the implementation of innovation, is simple yet profound. If educational designers, developers or innovators hope to create materials which are more efficient and/or effective than currently used materials, then they must also invest in preparing users (students)

BEST COPY AVAILABLE

for that change. Simply said, even 'perfect' material is inferior if a student does not use it effectively.

References

Berg, T. & Brouwer, W. (1991). Teacher awareness of student alternate conceptions about rotational motion and gravity. Journal of Research in Science Teaching, 28, *1*, 3-18.

Confrey, J. (1990). A review of the research on student conceptions in mathematics, science, and programming. In C.B. Cazden (ed), Review of Research in Education, Volume 16, 1990. Washington, D.C.: American Educational Research Association.

D'Amour, G. (1979). Problem-solving strategies and the epistemology of science. In J. Lochhead and J. Clement (eds), Cognitive process instruction: Research on teaching thinking skills. Philadelphia: Franklin Institute Press.

Festinger, L. (1957). A theory of cognitive dissonance. Evanston, Ill: Row Peterson.

Gilbert, J. & Watts, D. (1983). Concepts, misconceptions and alternative conceptions: changing perceptions in science education. Studies in Science Education, 10, 61-98.

Hewson, P.W. (1980). Learning and teaching science. South African Journal of Science, 76, 397-403.

Hodson, D. (1990). A critical look at practical work in school Science. School Science Review, 70, 33-40.

Kirschner P.A. & Meester, M.A.M. (1988). The laboratory in higher science education: Problems, premises and objectives. Higher Education, 17, 81-88.

Shuell, T. (1987). Cognitive psychology and conceptual change: Implications for teaching science. Science Education, 71, *2*, 239-250.

Suchman, J.R. (1966). A model for the analysis of inquiry. In J. Klausmeier and C.W. Harris (eds.) Analysis of concept learning. New York: Academic Press.

Theobald, D.W. (1968). An introduction to the philosophy of science. London: Methuen.

VIPS - VISUAL INFORMATION PROCESSING IN STUDY TASKS

Hans Stolk, Centre for Educational Technology and Innovation, Open University, Heerlen, The Netherlands

Introduction

When we observe a person studying visually presented learning material, we see a visual scanning process in which the eyes jump from one location to another. Early research by Yarbus (1967) already showed that the individual scanning behaviour, such as the scan-path, is strongly influenced by the learning objectives given to the subject and by subject predisposition in general.

Eye movement recording can be used as a technique for the acquisition of qualitative and quantitative data of the visual scanning process of learners that are studying specific learning material. Learners vary in the amount of time they spend

in text and picture areas of material, in the number and order of switches from text to picture and vice versa and in general reading behaviour. One of the questions which this research addresses is whether there are characteristics in the scanning process, such as the number of regressions and the number of text-pictures switches, that can be related to the learner's expectation and prior knowledge state. For the study tasks, learning material is presented on a video display unit consisting of subsequent screens with text and pictures. The topics of the study tasks vary and are taken from mathematics, law and other disciplines.

Apparatus
The eye movement recordings were carried out using a video-based eye tracker, which calculates the point of regard of a subject every 20 ms. The eye movement patterns are superimposed on a video image of the stimulus field, which allows for easy video recording and the numerical data are sent to a remote computer system for later analysis. For eye movement recording several techniques can be used. The system we developed, uses the corneal reflection/pupil centre method and calculates the point of regard with an accuracy of about 1 degree of arc.
The system has a low subject awareness, which means that it is not necessary to fixate the head of the subject and that no equipment is attached to the head. This way better ecological validity of the experiments is achieved. The 20 ms temporal resolution is sufficient for this type of research, as fixation durations are in the order of 200 ms.

Method
Subjects performed three subsequent study tasks in which they were asked to study stimulus material that was presented to them on a 20 inch video monitor. Subjects controlled the pacing of the presentation. The stimulus material consisted of a legal text about private law, a set of graphs and a mathematical text about the proof that the sum of the angles in a triangle is 180 degrees. The first task consisted of text only, the second was picture only and the last was a text-picture combination.
While the subjects were performing the tasks their eye movements as well as their general behaviour were recorded on video tape. After completion of the tasks, the subjects were asked to fill in a test which measured to what extent they understood the texts and were able to recall them.

Text/picture processing
In many situations people are in a process of learning from text and/or pictures. Reading in such a situation is not the straightforward processing of a piece of text. Often jumps from text to picture and vice versa will occur. Some sort of integrative process will take place, in which information from both sources is integrated into a knowledge structure inside the head of the reader. Jumps between text and picture will occur in order to verify information obtained from one source with information obtained from the other or to add information from one source to information from the other. Our knowledge of the reading process is characterized by the fact that we know much more of the lower order processes like letter and word recognition than of the higher order processes like discourse processing and learning [Rayner & Pollatsek, 1989]. This can be explained by the assumption that the lower order

processes are determined more directly by the serial aspects of saccadic eye movements and that in the higher order processes more parallellism is induced. At a lower level seperate words are put into the system in a sequential fashion. At a higher level words are integrated in parallel into a larger knowledge structure. Scanning patterns in picture viewing show ressemblances but also great differences to reading. On evolutionary grounds, we can say that writing developed from drawing simple sketches (Földes-Papp, 1966; Klix, 1980) and reading can thus be assumed to be a highly skilled form of picture viewing. One major difference is the role of instruction. Picture viewing usually starts with a fast, rough scanning of the main properties of the picture, which takes little time. What follows depends on the instruction given to the subject, but seems to be related to the content of the picture. The eyes tends to fixate on regions that are semantically rich (Buswell, 1935; Yarbus, 1967; Kolers, 1973; Loftus, 1974), although it is not always clear what is meant exactly with the notion of semantic richness. Noton and Stark (1971) reported typical scanning patterns in picture viewing that were less dependent on the picture content, the so-called 'scan-paths'.

Results

In order to get a better understanding of the relation between the content of the stimulus material and the nature of the eye movements we have tried to describe them using a more or less formal notation. We started with a description of the stimulus material. The description is based upon the descriptional system of Kintsch and Van Dijk (1978). In contrast with Kintsch and Van Dijk our description does not split up elements like [two points] into seperate elements. Table 1 shows the representation of two stimulus fields that belong to the mathematical proof.

Table 1

Description of the stimulus material (text).

T1		
1	[fact: path, shortest]	The shortest path
2	[location: between, two points, 1]	between two points is
3	[statement: is, 2, straight line]	a straight line.
T4		
1	[condition: when, 4]	When a straight line
2	[fact: line, straight]	intersects two parallel
3	[fact: two lines, parallel]	lines, the angles
4	[operation: intersect, 2, 3]	opposite to each other
5	[fact: angles, opposite of each other]	are equal.
6	[conditional statement: equal, 5]	

Similar descriptions were made for the other study tasks. The fixation sequences of the subjects were denoted in a protocol in a similar way, which allowed for easy comparisson. Table 2 shows an example of such a protocol.

Table 2
Sample protocol of a subject on T1-P1

1	[`The .. line.']	The shortest path
2	[picture]	between two points is
3	[`The'] a	straight line.
4	[end]	

The description in table 2 indicates that the subject read the text from the beginning till the end, then jumped to the picture and finally refixated the first word `The'. After this sequence the subject proceeded with the next paragraph.

Discussion
Analysis of the protocols for different subjects shows that eye movement patterns are influenced by the length and complexity of the material, but also depend on the characteristics of the subject. The location from which first jumps are made is usually determined by the position of periods or commas. This means that these first jumps mainly depend on punctuation and less on content.
A subject that makes numerous jumps from one part of the text to another or from text to picture vice versa tends to do so for different stimulus fields. Subjects that are unfamiliar with the topic of the study task show more refixations of earlier parts of the text and the total time these subjects spend on the task tends to be longer. Some of these subjects show more `jumpy' patterns. When the stimulus material is very simple, individual differences are relatively small. When the stimulus material is getting more complex, as was the case in the graphs, or when the text is getting longer, as was the case in the legal text, individual differences are numerous.

References

Buswell, G.T. (1935). How people look at pictures. University of Chicago Press: Chicago.

Földes-Papp, K. (1966). Von Felsbild zum Alfabet. Belser: Stuttgart.

Kintsch, W., & Van Dijk, T.A. (1978). Towards a model of text comprehension and production. Psychological Review, 85, 363-394.

Klix, F. (1980). Erwachendes Denken. Eine Entwicklungsgeschichte ser menschlichen Intelligenz. VEB Deutscher Verlag der Wissenschaften: Berlin

Kolers, P.A. (1973). Some models of representation, in Pliner, P., Krames, L. & Alloway, T. (eds.) Communication and affect: language and thought. New York: London.

Levin, J.R. (1981). On functions of pictures in prose, in Pirozzolo & Wittrock (eds.) Neuropsychological and cognitive processes in reading. Academic Press: New York.

Loftus, G.R. (1974). Acquisition in information from rapidly presented verbal and non-verbal stimuli, Memory and cognition, 2, 545-548.

Noton, D., & Stark, L. (1971). Eye movements and visual perception. Scientific American, 224(6), 34-43.

Rayner, K., & Pollatsek, A. (1989). The psychology of reading. Prentice Hall: Englewood Cliffs, New Jersey.

Yarbus, A.L. (1967). Eye movements and vision. Plenum Press: New York.

IS THE PROVISION OF EMBEDDED SUPPORT DEVICES, IN STUDY MATERIALS, IN CONFLICT WITH THE INDIVIDUAL LEARNING PROCESS?

M.M.A. Valcke, R.L. Martens, P. Poelmans, Centre for Educational Technology and Innovation (OTIC), & M.M. Daal, Centre for Educational Production, Open University, Heerlen, The Netherlands

Introduction

The Open university is an educational institution that delivers education at a distance. The structure and content of studies at the Open university are organized in a very open ended way (e.g. there are no starting conditions). In this particular educational setting students are provided with highly elaborated study materials. These materials are mainly text-based. A large set of embedded (in-text) support devices (e.g. adjunct questions, advance organizers, schemes, etc.) are used to support the individual learners during their learning process. The S.I.G.-project (Support by Individual Guidance) of researchers of the Centre for Educational Technology and Innovation (OTIC) focuses on the evaluation of these in-built support devices.

Functions and effects of embedded support devices: Key questions?

Embedded support devices (ESD) comprise a large set of structuring, activating and motivating elements in study materials. Analysis of Open university courses reveals that up to 44 different types of ESD are used. Among them, the most commonly known are : pretests, objectives, questions (pre- and post-), tasks, advance organizers, schemes (concept maps), tests, feedback, repeat units, typography, text structure, headings, illustrations, diagrams, etc.

These embedded support devices have been researched to a considerable extent in educational literature. But this research body and the research results have some shortcomings. First of all there are theoretical and methodological constraints : Most research can be described as effect-studies with a limited theoretical base focusing on potential functions of the ESD. Research from a more recent vintage questions this - reductionist and behaviouristic - approach (Winn, 1990). The question is put forward whether external regulation of learning is not conflicting with the individual learning process needs. The importance of internal regulation is put forward and the importance of individual differences. An additional, but important, constraint in the context of our study is the questionable ecological validity of most studies. Next to the fact that rather specific ESD are studied in artificial contexts, studies have hardly been set up in a distance education setting.

Next to the key question about the potential function and effects of ESD in a distance education setting, there is the other key question of the actual study about the actual use of the ESD by the students. Clyde and others state for instance : "little systematic knowledge is available concerning the ways in which students use distance teaching materials provided to them" (1983, p.4). Here we enter the issue introduced by the title of this conference contribution : Is the provision of

embedded support devices, in study materials, in conflict with the characteristics of the individual learning process ? In other words, is the assumption empirically grounded that ESD meet the demands and needs of the individual learner in the specific open and distance education setting ?

General outline of the study
To evaluate this basic assumption the study was split up into four main subparts :

Explorative phase
During this first phase literature was analyzed to gather an extensive overview of the potential functions and effects of a large set of ESD (Valcke, Dochy & Daal, 1991).

Analysis of Ou-course materials
The result of phase 1 was used to develop an instrument to detect and evaluate the extent to which ESD are used in Open university (Ou) study materials. This resulted in the development of a manual implying the use of checklists. Next, this instrument was used to analyze in detail the course "Introduction to Governmental Law". This try-out was discussed with law experts of the Ou-law department (Martens, Daal, & Valcke, 1991).

Theory development
To direct our research a theoretical model was elaborated to describe and explain the interrelation between the large variety of the ESD and the learning process (Valcke & Martens, 1992). This theoretical base builds on a cognitive psychological framework and integrates the three-component theory of Sternberg (1986), a model of the reading process, the impact of individual variables and features of the task environment.

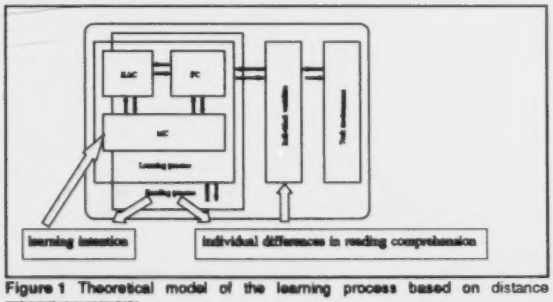


Figure 1 Theoretical model of the learning process based on distance education materials

The reading process is incorporated into the model because most Ou-study materials are still text-based. Of importance, for the focus of this article, is the relation with individual variables. The model in figure 1 was used to state the theoretical function and effects of the specific ESD (44) in our study.

The empirical study

Making use of the output of the three preliminary phases, an empirical study was set up.

A representative sample of 25 law-students, who enroled for the final test of the Ou-course "Introduction to Governmental Law" was interviewed, making use of a structured interview.

In the first part of this interview the general study interest, approach and strategy was discussed. Next the discussion concentrated on issues related to studying law related study content. In the last part of the interview, the interview concentrated on the specific study process of the course "Introduction to Governmental Law". In this part of the discussion, the interviewer and the student read through a complete learning unit. This learning unit was dissected by the researchers beforehand and all embedded support devices were highlighted. Student were invited to indicate what, how, when and why they used each of the specific embedded support devices. This open question was - after the answer of the student - extended with closed questions to check whether the students agreed with the functions and effects of the specific ESD as derived from the theoretical model developed by the researchers. The data obtained reflected in this way spontaneous reflections of the students about functions and effects of specific ESD and ideas about their agreement with the hypothetical theoretical functions and effects of ESD as postulated by the theory.

The tape-recorded interviews were analyzed following a checklist. The information was coded along up to 500 variables. Next, clusters of these variables have been analyzed using log-linear and hierarchical log-linear analysis techniques. In this way, the analysis of data is helpful to find conclusive information about the inhibitive or facilitative effect of embedded support devices in relation to the individual learning process. In other words, we check whether the results indicate that there is inference of embedded support devices with internal regulation needs during learning. The interview part of the empirical study was extended with a confirmatory check with tutors of the Ou-study centres.

Results

In the context of this short research report it is not possible to give a detailed account of the research results. A report is available on request (Valcke, M., Daal, M., Poelmans, P. & Martens, R., 1992). An extensive overview of the results of this study will be presented during the conference in view of the key question about the link between individual student needs and the 'standard' provision of ESD in study materials. In general these result indicate the following :

- The theoretical expectations about the functions and effects of ESD are not always reflected in the actual use of the ESD by the students.
- Student profiles in a distance education setting do differ to this extent that the choices made by course developers during the design of the study materials does not meet the specific needs and demands of individual students.

The results of this study set forth new demands for the development of study materials. At the OTIC a research project is set up to realize such a new approach. A flexible and automated learning environment is being conceptualized and developed in which student can choose more freely what kind of support devices, when, where, etc they want to be present in the study materials.

References

Clyde, A., Crowther, H., Patching, W., Putt, I., & Store, R. (1983). How students use distance teaching materials: an institutional study. Distance Education, 4, 1, 4-26.

Martens, R., Daal, M., & Valcke, M. (1991). Het gebruik van ingebouwde begeleidingscomponenten in de cursus "Inleiding tot het Bestuursrecht". Report U91/23.400. Heerlen : Open university.

Valcke, M., Daal, M., Martens, R., & Dochy, F. (1991). Opsporen en Begeleiden van Ingebouwde Begeleidingscomponenten - Handleiding. OTIC research report 42. Heerlen : Open university.

Valcke, M., Dochy, F., & Daal, M. (1991) : Functions and Effects of support in Open Learning Systems. OTIC research report 29. Heerlen : Open university.

Valcke, M. & Martens, R. (1992). The actual use of embedded support devices in self-study materials by students in a distance education setting - Theoretical base and research approaches. OTIC research report 53. Heerlen : Open university.

Valcke, M., Daal, M., Poelmans, P., & Martens, R. (1992). The actual use of embedded support devices in self-study materials by students in a distance education setting - Research Results. OTIC research report 54. Heerlen : Open university.

Winn, W. (1990). Some implications of cognitive theory for instructional design. Instructional Science, 19, 53-69.

THE RELATION BETWEEN COGNITIVE AND METACOGNITIVE SKILLS

Bernadette H.A.M. van Hout-Wolters, Graduate School of Teaching and Learning, University of Amsterdam, Amsterdam, & Jos J. Beishuizen, Department of Educational Psychology, Leiden University, Leiden, The Netherlands

Metacognitive skills have been identified as being important to cognitive performance in several domains (Campione, Brown, & Ferrara, 1982; Glaser, 1990). However, there are several relevant questions that remain to be answered about the specific role that metacognitive knowledge plays in cognitive performance. A general question concerns the definition and role of metacognition in relation to other factors identified in the literature as affecting cognitive performance and learning (such as intelligence, prior domain-specific knowledge and cognitive style). The first paper by Marianne Elshout-Mohr presents a general model specifying the nature of these relationships in an educational context. Each of the other three papers will look deeper into one of these relationships, both from a theoretical and an empirical point of view.

The paper by Veenman, Elshout and Busato addresses the issue of whether metacognitive skill is a general part of the cognitive apparatus. Specifically, the paper examines to what extent metacognition represents general intellectual abilities, and concurrently, to what extent metacognitive performance is malleable through specific interventions. A second issue, discussed in the paper of Samarapungavan and Milikowski, concerns the embeddedness of metacognitive knowledge in domain-specific knowledge. In this paper the nature of metacognitive knowledge is explored in terms of the declarative vs. procedural distinction. The last paper by Stoutjesdijk and Beishuizen also examines the role of metacognitive vs. domain knowledge in reading comprehension. Further, it also investigates the effect of idiosyncratical cognitive styles on reading comprehension in relation to metacognition.

In presenting both theoretical and empirical analyses of the relation between metacognition and cognitive skills, the four papers draw on different settings, task contexts, and domains. The Elshout-Mohr study takes a broad educational setting as the bases for its theoretical model. The Veenman et al. paper uses a discovery microworld in the domain of electricity to examine learning in a laboratory context. The Samarapungavan and Milikowski paper also examines scientific reasoning, however in theory-choice tasks in astronomy and chemistry. Finally, Stoutjesdijk and Beishuizen investigate reading comprehension from an electronic text and

their domain is psychology. This variation in tasks and domains may result in differing emphases on certain characteristics of metacognition. Because a general model of the relation between metacognitive and cognitive skills should incorporate these variational consequences, they are a point of discussion as well.

THE RELATION BETWEEN COGNITIVE SKILL AND METACOGNITION

Marianne Elshout-Mohr, SCO, University of Amsterdam, Amsterdam, The Netherlands

In this paper research on the relation between cognitive skill and metacognition is reviewed in order to better understand and explain factors that link cognitive and metacognitive growth. Insight into these factors is prerequisite for developing more effective ways to promote the acquisition of cognitive skills and metacognition. A major source of information about acquisition of knowledge based competencies and metacognition is research on novices and experts. This research seems to suggest that cognitive and metacognitive growth evolve concurrently, hand in hand so to say. Nothing much is said however, about factors that foster transitions from one stage of cognitive skill or one level of metacognitive awareness to the next.

One can speculate that characteristics of the learner determine what skills or metacognitions develop first and what newly acquired skill or awareness triggers the next. Among these are intelligence, prior knowledge, domain-general strategies, dispositions ("habits of the mind"), and already acquired skills in metacognitive reflection and control. Research on these factors confirms that, in general, cognitive skill and metacognition are mutually supportive.

Comprehension monitoring (a metacognitive skill), for instance, in general facilitates knowledge use and acquisition. This simple relationship is affected however, by factors like amount of prior knowledge and dispositions (e.g. orientation toward learning). Students with a performance orientation toward learning, for instance, are apt to withdraw when confronted with difficulties in understanding study material. For these students awareness of occasional failure to construct meaning and attain understanding can be a detrimental instead of a helpful component of the learning process.

The review of research findings promotes understanding of the relationship between cognitive skills and metacognition in performance and learning under different conditions.

METACOGNITIVE MEDIATION IN LEARNING WITH COMPUTER-BASED SIMULATIONS.

Marcel V.J. Veenman, Jan J. Elshout, and Vittorio V. Busato, Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

This paper elaborates on the issue of by relation between intellectual and metacognitive skills. In two extreme models representing this relationship, metacognitive skill is regarded either as just another manifestation of intelligence, or as an independent predictor of learning. Previous work of Veenman and Elshout (1990, 1991) however, provided evidence in favor of a mixed model. Metacognitive skill appeared to be related to intellectual ability, but also had a unique predictive value to learning. If metacognitive skill is less a matter of acquired technique related to intelligence than, for instance, of good habit or invested effort, then metacognitive instructions may raise the level of metacognitive activity as well as enhance learning performances.
The objective of the present study is to compare metacognitive mediation to unguided discovery learning in a computer simulation environment. High and low intelligent students worked in either a metacognitive mediated (MM) or unguided discovery (UD) environment for learning principles of electricity. Analyses of thinking-aloud protocols showed that MM subjects exhibited a better working method (reflecting metacognitive skill) than UD subjects. MM subjects also outperformed UD subjects on several posttest learning measures. However, metacognitive mediation did not have long-lasting effects on learning outcomes of low intelligent students. The results will be discussed in terms of their theoretical implications for the relation between intellectual ability and metacognitive skills.

THE RELATION BETWEEN METACOGNITIVE AND DOMAIN KNOWLEDGE IN THEORY-CHOICE CONTEXTS.

Ala Samarapungavan and Marisca Milikowski, Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

This paper examines two theoretical issues that concern the relation between metacognitive and domain specific knowledge. The first is the domain generality of metacognitive knowledge. We think that metacognitive control processes operate in a heuristic manner in interaction with domain specific beliefs. The implementation of metacognitive monitoring strategies such as the search for empirical anomalies or internal inconsistencies with regard to one's hypotheses is semantically mediated by one's domain knowledge. We suggest that metacognitive knowledge grows in fluency and efficiency through exercise in numerous and varied task-specific contexts.
The second issue is the question of conscious access to and verbalizability of metacognitive knowledge. We challenge the view of metacognitive knowledge as

"stable and statable knowledge" about cognitive activities (Brewer & Samarapungavan, 1991; Samarapungavan, in press). In this sense we depart from researchers like Kuhn (1989) who claim that metacognitive knowledge must be consciously statable. We argue that metacognitive knowledge may take a purely procedural form that is not normally accessible to consciousness and may not be "statable." Conversely, people may be able to verbalise certain kinds of metacognitive rules of performance without being able to implement them in task contexts.

READING STRATEGIES FOR LARGE EXPOSITORY TEXTS

Evelien Stoutjesdijk and Jos J. Beishuizen, Department of Educational Psychology, Leiden University, Leiden, The Netherlands

Prior research has highlighted the importance of both the readers' domain knowledge and their metacognitive strategies for effective reading comprehension. However, most models of reading comprehension are based on an analysis of what readers do with short texts. This study examines the role of prior domain knowledge and reading strategies on reading comprehension using a long (20,000 words) expository text on memory.

The text was presented electronically on computer screen in a series of "frames". Subjects could access the text in a frame by clicking on its concept icon. The text was accompanied by a map of concepts that graphically represented the relations between concepts contained in the text frames and also gave short descriptions of each concept.

The subjects were given a reading task which required them to locate and study a certain sub-section of information in the total text. To do so subjects had to locate and read 22 text frames out of the 72 frames available. There were 3 experimental conditions. In the first condition subjects received an advance organizer with an overview of the text's contents, thus enhancing their domain knowledge at the onset of reading. In the second condition subjects received strategic instructions on the use of the concept map and were advised to take notes during reading. The third condition was a control condition in which the subjects were only given a description of their reading task. All subjects were pretested on the Vermunt learning styles test to identify if they had an orientation towards self-regulation or external regulation. It turned out that subjects did not profit from either an advance organizer or strategic advice as far as their ability to locate relevant information is concerned. Lack of expertise in the domain of memory has probably caused the subjects to consider all information relevant in relation to the reading assignment. However, irrespective of the experimental treatments, different reading strategies were detectable in the protocols. Some subjects consistently used the available facilities for orientation on text frames before assessing and reading them, whereas other subjects read without orienting activities. Another dimension, on which reading strategies differed was the way in which subjects determined a reading path. Some students used navigation tools to select the next frame to assess, while other subjects followed the suggested ordering of items on the map of concepts.

References

Brewer, W.F., & Samarapungavan, A. (1991). Children's theories vs. scientific theories: Differences in reasoning of differences in knowledge. In R.R. Hoffman & D.S. Palermo (Eds.). Cognition and the symbolic processes. Hillsdale, NJ: Erlbaum.

Campione, J.C., Brown, A.L., & Ferrara, R.A. (1982). Mental retardation and intelligence. In R.J. Sternberg (Ed.), Handbook of human intelligence (pp. 392-490). Cambridge: Cambridge University Press.

Glaser, R. (1990). The reemergence of learning theory within instructional research. American Psychologist, 45, 29-39.

Kuhn, D., Amstel, I., & O'Laughlin, M. (1989). The development of scientific thinking skills. Orlando, FL: University of Florida Press.

Samarapungavan, A. (in press). Children's judgements in theory choice tasks: Scientific rationality in childhood. Cognition.

Veenman, M.V.J., & Elshout, J.J. (1990). De meerwaarde van een goede probleemaanpak. (The surplus value of an effective working method). Tijdschrift voor Onderwijsresearch, 15, 337-314.

Veenman, M.V.J., & Elshout, J.J. (1991). Intellectual ability and working method as predictors of novice learning. Learning and Instruction, 1, 303-317.

AUTHORS INDEX

Name	Chapter	Page
Aalsvoort, G. M. van der	3	375
Aarnoutse, C. A. J.	3	295
Aarnoutse, C. A. J.	5	719
Achtenhagen, F.	7	945, 999, 1035, 1037
Ackerman, T. A.	5	705
Admiraal, W.	6	823
Aguilera, A.	6	908, 911
Aho, L.	2	172
Akker, J. van den	2	99, 153, 189
Akkök, F.	5	731
Allegrezza-Carvoyeur, L. S.	9	1251
Alvarez, P.	5	668
Amelsvoort, J. van	3	396
Anderson, C.	8	1130
Ansellem, M.	3	320
Antoli, B.	6	894
Antunes, H.	9	1243
Ardaç, D.	9	1268
Askar, P.	3	258
Askar, P.	5	731
Assink, E. M. H.	3	285
Assink, E. M. H.	4	644
Atkins, M. J.	4	531, 651
Babanin, L.	4	643
Barendse, P.	8	1199
Barker, P.	4	592
Barrett, E.	6	854
Barrows, H. S.	7	954
Barton, L.	6	854
Bassey, M.	2	213
Bates, I.	9	1241
Becher, T.	8	1063
Behrens, M.	7	1001
Beishuizen, J. J.	3	432, 436
Beishuizen, J. J.	4	633
Benins, A.	6	904
Bennett, N.	3	400
Bennett, N.	6	939, 941
Berberoglu, G.	5	708, 748
Berg, E. van den	2	146
Berg, G. van den	8	1158
Berg, J. van den	7	1020
Berg, R. van den	4	521
Bergen, J. van	5	786
Bergen, Th.	3	396, 396
Bergen, Th.	6	931
Berger, M. P. F.	5	692
Berger, N.	3	387
Bergh, H. van den	2	108
Bergmans, A.	3	295
Berings, I.	9	1286
Bestebreurtje, R.	4	584
Bijtel, J. A. L. van	6	796
Black-Branch, J. L.	9	1276
Blackman, S. J.	4	622
Blom, S.	1	79
Boei, F.	6	931
Boekaerts, M.	3	302, 360, 395
Boekenoogen, M. S.	7	994
Boekhorst, A. K.	4	619
Boekkooi-Timminga, E.	5	698
Bock, R. D.	5	662, 753
Bögels, S. M.	3	245
Bokhove, J.	5	783
Bollen, R.	1	71
Bolzoni, M.	4	569
Bonekamp, L. W. F.	4	538
Boon, J.	7	968
Börger, M. A.	6	924
Bosch, H. van den	8	1201
Bosch, L. van den	5	786
Bos, K. Tj.	1	63
Bos, K. Tj.	6	862
Bosker, R. J.	1	19, 86
Bosker, R. J.	9	1215
Bosman, Y.	7	1004
Bostock, S. J.	4	502
Bossink, C.	6	922
Bouhuijs, P. A. J.	8	1187, 1188
Braaksma, J.	4	577
Brandsma, H. P.	1	8, 69
Brandsma, H. P.	6	844
Brandsma, T. F.	7	1040, 1053
Bratko, D.	5	735
Brekelmans, M.	6	935
Breuer, K.	3	415
Brinke, D. ten	5	774
Brinke, J. S. ten	6	820, 932
Bristow, S.	3	405
Bristow, S.	9	1257
Brouwer, K.	7	964

Brown, A. 7 1001
Brown, A. 8 1168
Bruggencate, G. C. ten 8 1107
Bruijn, E. de 7 1015
Brummelhuis, A. ten 2 206
Bruinsma, C. 6 837
Büchel, F. P. 3 407
Büchel, P. 3 407, 409
Burton, R. M. 1 94
Busato, V. V. 3 435
Cachapuz, A. 6 816
Calderheads, J. C. 6 878
Cardelle-Elawar, M. 3 371
Cardelle-Elawar, M. 6 829
Carleer, G. J. 2 186
Carleer, G. J. 4 434, 631
Carré, C. 6 939, 940
Cate, Th. J. ten 6 796
Cate, Th. J. ten 8 1141
Cavendish, S. 9 1221
Cerveró, F. 6 894
Chartier, D. 3 329
Christiaans, H. H. C. M. 3 401
Claessens, M. P. J. G. 8 1196
Clarke, S. 7 1010
Clement, M. 6 885
Cohen, R. 4 587
Collis, B. A. 4 439, 534, 631
Collins, D. 4 502
Commins, B. 4 525
Confalonieri, E. 9 1245
Connelly, F. M. 2 213
Corporaal, A. H. 6 931
Corte, E. De 3 351, 393, 398, 417
Cox, M. 4 631
Creemers, B. P. M. 1 74, 82
Cremers, L. 4 545
Cremers, L. 6 862
Créton, H. 6 935
Cuijk, H. van 9 1286
Czerniawska, E. 3 345
Daal, M. M. 3 429
Daal, V. H. P. van 3 293
Daly, P. 1 60
Dam, G. T 9 1218
DeAyala, R. 5 686
Dekkers, H. 9 1226
Demirel, M. 3 342
Desforges, C. 3 405
Desforges, C. 9 1257
Diana, I. De 4 595, 598
Diana, I. De 5 728
Diggele, J. van 5 783
Dijk, H. van 3 263
Dijk, W. van 3 354
Dijkstra, S. 3 223, 414
Dirkzwager, A. 4 634
Dochy, F. J. R. C. 3 398, 417
Doets, C. 7 1043
Dolmans, D. 8 1133
Donald, J. G. 8 1145
Dong, Z-P. 4 557
D'Onofrio, A. 8 1172
Doornekamp, G. 4 545, 605
Doughty, P. L. 4 625
Draper, J. 6 813
Driel, C. van 6 820
Driel, J. van 8 1189
Dronkers, J. 9 1271
Dumitrescu, A. 2 131
Dumitrescu, A. 6 856
Dunne, E. 6 939, 940
Dunne, R. 6 871
Eash, M. J. 2 169
Eck, E. v. 9 1218
Eberg, J. W. 2 111
Eberg, J. W. 7 994
Edelenbos, P. 1 69
Eeden, P. van den 2 100
Eggen, T. 5 786
Ehrlich, D. 4 572
Eijkelhof, H. M. C. 3 256
Eldering, L. 9 1282, 1283
Elen, J. 8 1178
Elliott, R. G. 6 834
Elshout, J. J. 3 310, 435, 461
Elshout-Mohr, M. 3 434
Elst, G. van 7 977
Ely, D. P. 4 480
Engelen, R. 5 786
Engelhard, G. 5 741
Erder, A. M. 3 342

Ersoy, Y. 4 483
Ersoy, Y. 6 896
Etelpelto, A. 7 951
Evans, D. 1 55
Fátima Choráo C. Sanches, M. de 3 269
Feijen, C. J. 7 1029
Fergusson-Hessler, M. 3 232
Fernandez Benitez, V. 6 858
Feteris, A. 7 1017
Finegold, M. 4 457
Fischer, D. 6 936
Fitz-Gibbon, C. T. 1 16
Fleming, R. 6 904
Forcheri, P. 3 323
Fox, M. 8 1172
Fraser, H. 6 813
Fraser, B. 6 936
Free, E. L. 8 1114
Frey, K. 2 197
Fricke, R. 4 617
Frijns, P. 8 1193, 1199
Furlong, J. 6 854
Gaag, F. J. van der 4 613
Geban, O. 3 258
Geest, T. M. van der 2 187
Gelderblom, A. 7 980
Geus, W. de 3 298
Geus, W. de 5 744
Gielen, I. 3 351
Giesbers, J. 1 90
Giesbers, J. 6 931
Gijselaers, W. H. 8 1133, 1202
Gillijns, P. 5 786
Glas, K. 5 774
Glendinning, A. 9 1254
Glopper, K. de 3 298
Gmelich Meijling, M. R. 4 490, 613
Goldstein, H. 1 11
Goldstein, H. 5 751
Gómez-Crespo, M. A. 3 381
Gorsky, P. 4 457
Graaff, E. de 8 1187, 1188, 1193, 1208
Grace, G. 6 828
Gravemeijer, K. 2 103
Gray, J. 1 40
Groen de Jong, C. 3 364
Groot, M. de 3 375
Groot, T. 4 619
Groppo, M. 9 1245
Gruber, H. 7 1038
Grunefeld, H. 8 1069
Guillén, J. 3 390
Gulmans, J. 4 521
Haan, D. M. de 1 75
Haapala. A. 2 172
Hacquebard, A. E. N. 2 165
Halász, G. 9 1275
Hambleton, R. K. 5 681
Hamers, J. 3 407
Hameyer, U. 2 99, 153
Hansen, K-H. 4 487
Harlen, W. 2 175
Harvard, G. 6 871
Heathcote, G. 6 793
Heitmann, G. 8 1126
Hell, J. G. van 4 461
Helmke, A. 3 398
Helsper, E. 4 640
Hendry, L. B. 9 1254
Henneken, O. W. H. 5 744
Hermans, H. 3 422
Herrlitz, W. 2 212
Heuvel-Panhuizen, M. van den 2 103
Heuvel-Panhuizen, M. van den 3 312
Hippe, Z. S. 4 465
Hoeben, W. Th. J. G. 2 210, 212
Hoeven-van Doornum, A. A. 9 1264
Hofman, R. H. 1 46
Holmberg, L. 5 717
Holsbrink-Engels, G. A. 3 243
Hoogstraten, J. 8 1077
Hopkins, D. 1 62
Hough. M. 6 846
Hout, J. F. M. J. van 8 1155
Houtveen, Th. 6 802
Hout-Wolters, B. H. A. M. van

	3	432
Hout-Wolters, B. H. A. M. van	6	791, 930
Hövels, B.	7	1020
Huijsman, Z. C.	8	1081, 1153
Hulst, A. van der	3	357
Hutchison, D.	5	756
Huteau, M.	3	329
Ibrahim, A. B.	1	16
Izard, J.	5	775
Jagt, P. van der	6	837
Jahn, H.	8	1123
Jakowicka, M.	3	348
James, C.	6	878
Jansen, E. P. W. A.	8	1073
Janssen, J.	5	783
Janssen, J.	9	1228
Janssen, P. J.	3	402
Janssen, T.	7	1041
Janssen Reinen, I.	2	202
Järvelä, S.	3	272
Jaspers, J. G. M.	4	560
Jeffcut, P.	8	1081
Jenkins, H.	1	55
Jochems, W. M. G.	8	1084, 1088, 1196, 206
Johnson, S.	1	13
Jong, F. P. C. M. de	3	396
Jong, J. de	5	721
Jong, J. A. de	7	971, 977
Jong, M.J. de	6	865
Jong, R. de	6	806
Jong, W. J. de	8	1088
Joosten, G.	9	1228
Jorge, M.	6	816
Jost, K. L.	4	625
Jung, M. H.	4	617
Jungbluth, P.	9	1264, 1284, 1311
Jurgens, H. J. H.	8	1155
Kaldeway, J.	3	337
Kamp, L. J. Th. van der	5	759
Kamp, M. van der	7	1041
Kanselaar, G.	2	218
Kanselaar, G.	4	560
Karstanje, P. N.	1	42
Karstanje, P. N.	7	1004
Kattenberg, G.	3	285
Kauppi, A.	6	905
Kessels, J.	7	1025
Kesteren, B. J. van	4	534
Keursten, P.	2	143
Kieviet, F.	6	919, 931
Kirschner, P. A.	3	417, 422
Kleintjes, F. G. M.	1	25
Kley, P. van der	9	1286
Kluvers, C.	3	396
Knuver, A. W. M.	1	8
Koetsier, C. P.	6	820
Kohlemainen, P.	8	1111
Kohn, D. I.	3	384
Kok, W. A. M.	3	306
Kok, W. A. M.	4	560
Kommers, P.	4	580
Koning, J. de	5	739
Koning, J. de	7	980
Kontogiannopoulou-Polydorides, G.	2	198
Koolstra, C.M.	5	759
Koopman, P.	8	1100
Kop, J. L.	9	1251
Koper, E. J. R.	4	634
Korthagen, F.	3	337
Korthagen, F.	6	823, 930, 932
Korthals, M.	8	1182
Kovacevic, P.	5	735
Kraemer, J. M.	5	783
Krause, B.	3	267
Kremers, E. J. J.	1	25
Krogt, F. J. van der	7	973
Krykorková, H.	3	249
Kuiper, W. A. J. M.	2	155
Kusá, D.	6	916
Ladhani, A-N.	4	595
Landucci, M.	2	117
Lang, M.	2	200
Langereis, M.	8	1077
Lautrey, J.	3	329
Leal Filho, W. D. S.	2	121
Lebedev, A. N.	5	734
Leemkuil, H. H.	3	325
Leenders, Y.	6	900
Leeuwe, J. F. J. van	5	719
Leiblum, M. D.	4	549
Lera, M.	6	908

Letiche, H. K.	8	1081, 1153
Levy, F.	3	320
Levy, J.	6	936
Lewis, B.	4	468
Lijnse, P. L.	3	256
Lima, A. de	6	880
Linden, E. van der	4	567
Linden, J. van der	4	644
Linden, W. van de	5	657
Lindström, B.	4	632
Lingefjärd, T.	4	647
Liverta Sempio, O.	9	1245
Loarer, E.	3	329
Lockwood, F.	5	660
Lodder, B. J. H.	8	1057
Lodder, B. J. H.	9	1235
Logar-Djuric, S.	5	735
Lombardini, H. E.	2	117
Loo, P. J. E. van de	8	1057
Lopez, A. M.	5	762
Love, J. G.	9	1254
Lowyck, J.	3	415
Lowyck, J.	6	840
Luyten, H.	1	74
Macready, G.B.	5	686
Main, R. G.	3	368
Maltseva, I.V.	5	734
Mandl, H.	7	1038
Manfred, E.	5	685
Mantovani, G.	4	569
Mª Poulussen van Balen, J.	6	858
Marinov, S.	5	711
Markinev, A. V.	5	734
Martens, R. L.	3	429
Martinez, R.	9	1293
Martinkova, V.	2	128
Martiny, U.	8	1060
Mason, G.	6	936
Mason, L.	3	333
Masters, G. N.	5	664
Matthijsen, C.	6	844
McGuigan, L.	2	209
McGuigan, L.	3	261
Meer, L. v. d.	8	1182
Meer, S. van der	8	1133
Meester, M. A. M.	3	422
Meesterberends-Harms, G. J.	7	958
Meijer, P. J.	3	281
Meijer, P. J.	4	563
Meijer, R. R	5	665
Mellink, E. C.	7	1004
Merriënboer, J. van	3	411
Menis, J.	2	213
Meulemann, H.	9	1273
Meyer, J. R.	6	849
Meyer, P.	6	900
Middelbeek, E.	3	422
Miles, S.	6	854
Milikowski, M.	3	310, 435
Miller, L.	4	609
Millett, A.	2	208
Min, F. B. M.	4	490, 613
Minnaert, A.	3	402
Mirande, M. J. A.	4	549
Moelands, F.	5	786
Moerkamp, T.	7	1044
Molenaar, I. W.	5	673
Molfino, M. T.	3	323
Mooij, T.	2	100
Moonen, J.	4	439, 633
Mora, J. G.	6	911
Mora, J. G.	9	1293
Morais, A. M.	9	1243
Morales, M.	5	762, 764
Moreno	5	764
Moskovkina, A. G.	9	1262
Mourik, T. G. C. van	3	245
Mulder, F.	2	165
Mulder, J. W.	8	1117
Mulder, L.	9	1312
Mulder, M.	7	945, 1021
Munn, P.	8	1093
Muraki, E.	5	662
Naylor, A.	6	914
Neber, H.	7	1039
Neut, I. v. d.	4	446
Neuwahl, N. M. E.	3	364
Nicaise, I.	9	1231
Niederer, R.	2	197
Niet, G.	4	613
Nieuwenhuis, A. F. M.	7	991, 1015
Nieuwenhuizen, C	3	302
Nieveen, N.	2	143
Nijhof, W. J.	7	945, 1033

Nguyen-Xuan, A.	3	320
Obel, B.	1	94
O'Brien, J. P.	6	914
O'Connor, J. E.	4	454
Oestermeier, U.	3	236
Ojen, Q. H. J. M. van	1	42
Ojen, Q. H. J. M. van	6	851
Ok, A.	6	809
Olsen, J.	2	184
Olsen, J.	4	609
Onna, B. van	7	960, 1029
Onstenk, J.	7	1046
Oort, F. J.	5	708, 748
Oosterhuis-Geers, J.	8	1065
Otten, R.	3	395
Oquist, O.	2	141
Oud, J. H. L.	5	719
Pan, H.	1	11
Pan, H.	5	751
Pavia, F.	6	894
Paykoç, F.	6	809
Peisa, S.	6	905
Pelgrims, G.	3	407, 410
Pelgrum, W. J.	1	75
Pelgrum, W. J.	2	195, 202
Perez-Eccheverria, M. P.	3	378
Petrikás, A.	1	31
Pieters, J. M.	6	791
Pieters, M.	2	115
Pijnenborgh, F.	3	250
Pilot, A.	4	577
Plomp, Tj.	2	193, 195, 202, 213
Plomp, Tj.	4	534
Pluijm, J. van der	9	1259
Podolskij, A. E.	8	1181
Poelmans, P.	3	429
Pool, I.	7	977
Poorthuis, G. M.T.	3	306
Popham, W. J.	5	713
Poschner, M. T.	8	1165
Pouw, C. L. M.	8	1114, 1182
Powney, J.	2	213
Powney, J.	8	1161
Pozo, J. I.	3	378, 381
Prucha, J.	2	125
Puchová, I.	3	276
Qualter, A.	2	209
Raban, B.	2	208
Raffe, D.	7	985
Ramaekers, G. W. M.	8	1057
Rasbash, J.	1	11
Rasbash, J.	5	751
Reezigt, G. J.	1	34
Reezigt, G. J.	9	1314
Reimerink, H. B.	4	613
Reitsma, P.	3	289, 298
Reitsma, P.	5	744
Reitsma, P.	6	874
Reisby, K.	2	135
Renkl, A.	7	1038
Renyier, M.	2	117
Resing, W. C. M.	3	387
Reynolds, D.	1	62
Reynolds, L.	4	572
Rezendes, G.	5	685
Riet, K. van 't	4	493
Rijns, Y.	8	1088
Riley, K. A.	1	58
Ris, Th. P. J.	4	615
Robijns, M. C.	9	1279
Roebuck, M.	4	529
Roeleveld, J.	9	1279
Rodriguez, D.	3	390
Rodriguez, R.	6	936
Roes, M. G.	2	148
Rojero, F. F.	2	117
Roode, F. de	5	774
Ropo, E.	2	133
Ropo, E.	6	826
Ruijs, G. P. C. M.	7	1050
Russell, T.	2	206, 209
Russell, T.	3	261
Sagdullaev, A. A.	9	1262
Sainsbury, M.	5	779
Saldaña, D.	3	390
Salonen, U.	2	172
Samarapungavan, A.	3	435
Sanders, M.	6	900
Sanz, A.	3	378
Saroyan, A.	8	1145
Schadé, E.	8	1141
Schagen, I.	5	756, 776
Schaick, P. van	4	513

Schaick Zillesen, P. G. van 4 490, 493, 613
Scheerens, J. 1 3, 80, 85, 86
Schmidt, H. 8 1133
Schmidt-Hackenberg, B. 7 989
Schmitt, M. 3 285
Schoenmaker, J. 4 468
Schott, F. 3 416
Schott, F. 8 1181
Schouten, A. 4 637
Schranhorst, U. 3 407, 411
Schreurs, M. L. J. J. 7 947
Schuurs, U. 2 108
Schuurs, U. 5 721
Sciarone, A. G. 3 281
Sciarone, A. G. 4 563
Scaratti, G. 9 1245
Seegers, G. 3 302, 360
Segers, S.R. 5 772
Sijtsma, K. 5 665
Sijtstra, J. 5 784
Simons, P. R. J. 3 395, 396, 411, 417
Simons, R-J. 3 223
Siregar, N. 6 867
Sheridan, B. E. 5 671
Shucksmith, J. 9 1254
Slettenhaar, H. K. 8 1117
Sliepen, S. E. 6 874
Sligte, H. W. 4 509
Smeets, D. 2 165
Smid, H. J. 8 1088
Smith, K. 5 772
Snippe, J. 2 151
Sola, I. 3 390
Spencer, E. P. 1 5
Spoon, P. 4 446
Squarcy, K. 5 735
Staa, A. van 7 968
Staessens, K. 6 891
Stam, H. van der 8 1120
Stanchev, I. 4 439
Stanley, S. J. 2 138
Steiner, G. 3 239
Stenström, M. L. 9 1248
Stoel, W. G. R 6 844
Stokking, K. M. 1 49
Stolk, H. 3 425
Stoutjesdijk, E. 3 436
Streumer, J. N. 7 945, 1009, 1017
Strum, J. 2 212
Struyf, E. 3 351
Sultana, R. G. 9 1238
Sumner, R. 8 1102
Taconis, R. 3 232
Tartwijk, J. van 6 935
Taylor, W. 6 813
Terekhina, A. 4 641
Tergan, S-O. 3 236
Terlouw, C. 8 1114, 1175
Terwel, J. 2 99, 100
Thompson, J. 4 525
Tiemessen, W. J. M. 8 1117
Tillema, H. H. 6 919
Timmer, E. 8 1091
Timmer, J. 4 544
Tobi, H. 5 695
Toluk, Z. 6 896
Tomic, W. 2 168
Trumper, R. 2 225
Tuijnman, A. 2 206
Tymms, P. B. 1 37
Ulbricht, K. R. G. 5 767
Underwood, G. 9 1221
Underwood, J. 9 1221
Valcke, M. M. A. 3 429
Valk, A. E. van der 3 256
Vandenberghe, R. 6 882, 885, 888, 891
Vedder, P. 9 1282, 1283
Veen, A. M. 9 1279
Veen, C. H. J. van 8 1081, 1153
Veen, J. T. van de 8 1107
Veen, R. van de 8 1091
Veen, W. van 4 446
Veenman, M. V. J. 3 435, 461
Veenman, S. 6 900
Veenstra-Strijland, T. 2 165
Vegt, A. L. van der 6 802
Velden, L. v. d. 9 1318
Velden, R. van de 9 1235
Velden, R. K. W. van der 8 1057
Vendrig, A. S. A. 7 971

Verhagen, P. W. 4 518, 584
Verloop, N. 6 926, 934
Vermeulen, C. 1 19
Vermeulen, C. 6 862
Vermeulen, M. 1 42
Vermeulen, M. 8 1096
Vermunt, J. D. H. M. 7 983
Vermunt, J. D. H. M. 8 1079
Verschaffel, E. 3 351
Versloot, A. M. 7 971
Verwey, A. 8 1088
Vila, L. 9 1293
Vinjé, M. P. 5 785
Vinke, D. 8 1084
Visscher, A. J. 1 63
Visser, H. J. 1 42
Vlas, H. 4 545
Voeten, M. J. M. 3 295
Voeten, M. J. M. 9 1264
Vogelzang, F. 4 446
Volman, M. 9 1218
Volpe, R. 6 904
Volman, C. 5 728
Voncken, E. 7 1004
Voogt, J. M. 2 184, 188
Vooijs, M. W. 5 759
Voronel, Y. 9 1289
Vos, H. 3 228
Vos, H. J. 5 725
Vos, M. 1 28
Vouten, M. J. M. 5 719
Vries, S. de 4 580
Vrugt, A. 8 1077
Vu, N. V. 7 954
Walker, D. F. 2 163
Walraven, M. 3 289
Walsh, A. 2 208
Wang, J. W. 4 541
Warmerdam, J. 7 973
Watson, D. M. 4 441
Weerden, J. van 5 783
Weide, M. G. 1 34
Weide, M. G. 9 1315
Werf, G. van der 9 1296
Wesselingh, A. 9 1215, 1286
Westbury, I. 2 182
Weston, P. 2 179
Wetterling, J. 4 553
Wieringen, F. van 1 88
Wiggers, J. A. 7 991
Wierstra, R. F. A. 2 159
Wijngaarden, K. van 8 1138
Wilbrink, B. 5 701
Wilbrink, B. 8 1149
Wignall, R. 1 53, 67
Willems, J. 8 1138
Wise, S. L. 5 689
White, J. 6 800
Whitty, G. 6 854
Witziers, B. 1 21
Woodhouse, G. 1 11
Woodhouse, G. 5 751
Woord, J. van der 8 1208
Woude, J. van der 4 506
Wubbels, Th. 6 820, 823, 930, 936
Yang, M. 1 11
Yang, M. 5 751
Yildiz, R. 4 651
Yontar, A. 3 253
Yrjönsuuri, R. 3 317
Zarevski, P. 5 735
Zeeuw, G. de 4 509
Zhang, J-P. 4 557, 602
Zhu, Z. 4 598
Zimowski, M. F. 5 753
Zinn, B. R. 3 384
Zinn, B. R. 9 1289
Zolingen, S. J. van 7 960
Zuidema, J. 4 640
Zuilichem, J. van 7 983
Zysno, P. V. 5 680
Zwart, W-J. 4 498
Zwietering, M. H. 4 493

22 - 25 JUNE 1992

BOOK OF SUMMARIES

EUROPEAN CONFERENCE ON EDUCATIONAL RESEARCH

UNIVERSITY OF TWENTE

ENSCHEDE - THE NETHERLANDS

BEST COPY AVAILABLE

EDITORS: TJ. PLOMP, J.M. PIETERS, A. FETERIS

European Conference on Educational Research

BOOK OF SUMMARIES

Volume 2

University of Twente
The Netherlands
June 22 - 25, 1992

461

COLOFON

Editors:
Tjeerd Plomp
Jules Pieters
Andries Feteris

Editorial assistance:
Harmen Abma
Jeroen Breman
Conny de Koning
Olivia Kramers
Renate Schraa

Cover:
Hanna Schnijder

Print:
Duoprint

University of Twente
Department of Education
P.O. Box 217
7500 AE Enschede

ISBN 90-365-0535-6

Copyright © 1992 by Department of Education, University of Twente

462

CONTENTS

Chapter 4
EDUCATIONAL TECHNOLOGY 437

Papers
Educational technology: Instrumentation and implementation, *J. Moonen, I. Stanchev, & B. Collis* 439
Case studies of classroom processes using geography simulations, *D.M. Watson* 441
How computers are used in instructional practice: Four case–studies in a Dutch secondary school, *W. Veen, F. Vogelzang, I. v.d. Neut, & P. Spoon* 446
Participant based formative evaluation of a collaborative technology project: Working with IBM and the United States Department of Education, *J.E. O'Connor* 454
A computer based instructional strategy for restructuring students' concepts of force, *P. Gorsky, & M. Finegold* 457
Using the computer as a help tool during learning by doing, *M.V.J. Veenman, J.J. Elshout, & J.G. v. Hell* 461
Experiences in professional education of operators of large industrial objects using knowledge-based system, *Z.S. Hippe* 465
Multi media courseware in a learning context: Perspectives on the roles of media and human factors, *B. Lewis, & J. Schoenmaker* 468
Current trends in educational technology: A methodology and findings, *D. P. Ely* 480
On the introduction of computers into the Turkish educational system, *Y. Ersoy* 483
Why teachers teach content and context area topics of computing, *K-H. Hansen* 487
Computer support of operator training based on an instruction theory about parallelism, *F.B.M. Min, P.G. van Schaick Zillesen, & M.R. Gmelich Meijling* 490
Computer support of process engineering education, *P.G. van Schaick Zillesen, M.H. Zwietering, & K. van 't Riet* 493
Instructional transaction theory applied to computer simulations, *W-J. Zwart* 498
A dynamical configurable remote interface for educational computer conferencing, *D. Collins, & S.J. Bostock* 502
TeleCAI: Computer aided instruction at a distance, *J. v.d Woude* 506
Increasing competence in open tele-learning environments, *G. de Zeeuw, & H.W. Sligte* 509
Structured electronic communication in computer-managed instruction, *P. v. Schaik* 513

Formal features as a design factor of video segments in interactive video programmes, *P.W. Verhagen* 516
A videodisc program for the acquisition of diagnostic skills in health education, *J. Gulmans, & R. van den Berg* 519
The efficiency and effectiveness of interactive video in skills training: The enterprise in higher education, *J. Thompson, . & B. Commins* 523
The application of interactive video to the professional development of teachers across Scotland, *M. Roebuck* 527
Are they thinking? Problem solving and learning theories in an interactive video applications, *M.J. Atkins* 529
Procedures for supporting the selection of courseware by secondary school teachers, *B.J. van Kesteren, G.J. Carleer, B.A. Collis, & Tj. Plomp* 532
Summative and formative evaluation of educational software in European countries, *L.W.F. Bonekamp* 536
A model of evaluating courseware, *J.W. Wang* 539
From computerliteracy towards information society literacy? *J. Timmer* 542
Conditions for effective implementation of computers in secondary schools in the Netherlands: The experimental schools project, *H. Vlas, G. Doornekamp, & L. Cremers* 545
Implementing university CAL: A top–down approach, *M.J.A. Mirande, & M.D. Leiblum* 549
Cost-effectiveness analysis of three alternative ways of media use training, *J. Wetterling* 553
The portability of EWB package to China, *Z-P. Dong, & J-P. Zhang* 557
The added value of new technology in computer assisted second language learning, *J.G.M. Jaspers, G. Kanselaar, & W.A.M. Kok* 560
How free should students be? A case from CALL: Computer Assisted Language Learning, *A.G. Sciarone, & P.J. Meijer* 563
Does feed back enhance computer assisted language learning? *E. v.d. Linden* 567
Analysis criteria for the evaluation of decision support systems (DSS) used by public educational agencies (PEA) in vocational guidance (VG), *G. Mantovani, & M. Bolzoni* 569
Multimedia in industry and education: A decision model for design, *L. Reynolds, & D. Ehrlich* 572
Training library-users in searching literature systematically with computer assisted learning: The fidelity of a simulation, *J.Braaksma, & A.Pilot* 577
Hypermedia for education: The control of information, *S. d. Vries, & P. Kommers* 580
Generic multimedia database: New didactics, *R. Bestebreurtje, & P.W. Verhagen* 584
The use of voice synthesizer in the discovery of the written language by very young children, *R. Cohen* 587
Electronic books and their potential for distance learning, *P. Barker* 592
Courseware engineering using knowledge-based performance modeling tools, *A-N. Ladhani, & I. d. Diana* 595

464

Designing "Share": An environment for classroom-based teaching and learning, *Z. Zhu, & I. d. Diana* 598
Interactive audio learning environment, *J-P. Zhang* 602
The valuation by students of the use of computers in education, *B.C. Doornekamp* 605
Classroom change: Is there a steering effect of computer technology? *J. Olson, & L. Miller* 609

Posters
Two examples of simulation environments for operator training based on parallel instruction, *F.B.M. Min, P.G. v. Schaick Zillesen, H.B. Reimerink, M.R. Gmelich Meijling* 613
Organisation of computer applications in dutch secondary education, *Th.P.J. Ris* 615
Interactive video in software training-learning with multimedia, *M.H. Jung, & R. Fricke* 617
Giraffe: A computer assisted instruction programme, *A.K. Boekhorst, & T. Groot* 619
Understanding technology in secondary school: Can the teacher stay ahead of the pupil? *S.J. Blackman* 622
The implementation of technology in the classroom: An examination of teachers beliefs, practice, and context in shaping meaning and adaptation, *K.L. Jost* 625

Symposia
'Technology- enriched school' projects, *G. Carleer, & B.Collis* 631
Didactics and educational computer software, *A. Dirkzwager* 634

Chapter 5
RESEARCH METHODES AND EVALUATION RESEARCH TECHNOLOGY 655

Papers
Advances in research methodology and evaluation research, *W. v.d. Linden* 657
Self recorded audio tape: An alternative Qualitative data collection method, *F. Lockwood* 660
Scoring function and generalized partial credit model, *E. Muraki, & R. Darrell Bock* 662
Constructing measurement variables with the Rasch partial credit model, *G.N. Masters* 664
A person fit function sensitive to item reliability and distance between items, *R.R. Meijer, & K. Sijtsma* 665
Rasch analysis as an instrument for measurement in the case of missing data, *P. Alvarez* 668
Evaluating the development of critical thinking in undergraduate nurses using Rash measurement, *B.E. Sheridan* 671
Non-parametric models for multicategory: Item response models reviewed and extended, *I.W.Molenaar* 673

465

A multidimensional polytomous latent trait model that explains the examinatees' choice of partially correct alternatives in multiple choice items, *H. Kelderman* 676
Scalogram analysis on polytomous items, *P.V. Zysno* 680
Measurement advances to address educational policy question, *R.K. Hambleton* 681
A multiple linear regression approach to placement at the united states coast guard academy, *E. Manfred, & G. Rezendes* 685
An adaptive testing model for assessing factors impeding skill acquisition, *G. B Macready, & R. DeAyala* 686
Self-adapted testing: Promise and implications, *S.L. Wise* 689
Designing of optimal samples for the estimation of parameters in item response models, *M.P.F. Berger* 692
Optimal sampling designs for group-level IRT models, *H. Tobi* 695
Adapting infeasible test construction models, *E. Boekkooi-Timminga* 698
Modelling the connection between individual behaviour and macro-level outputs: Understanding grade retention, drop-out and study-delay as system rigidits, *B. Wilbrink* 701
Determining the optimal measurement direction in a multidimensional latent space, *T.A. Ackerman* 705
Using restricted factor analysis with binary data for item bias detection and item analysis, *F.J. Oort, G. Berberoglu* 708
Physical dimensional theory as a method for analysis of psychological data, *S. Marinov* 711
Twenty-five years in the testing trenches: Lessons learned about educationally beneficial achievement test, *W.J. Popham* 713
Answer questions or tell the story? *L. Holmberg* 717
LISKAL: A pupil monitoring system based upon a new longitudinal measurement methodology, *J.H.L. Oud, J.F.J. van Leeuwe, C.A.J. Aarnoutse, & M.J.M. Voeten* 719
An adaptive test of Dutch as a second language, *U. Schuurs, & J. de Jong* 721
Simultaneous optimization of classification decisions followed by a mastery decision, *H. J. Vos* 725
A decision support system for the evaluation of courseware quality, *C. Volman, & I. d. Diana* 728
Adaption and standardization of the teacher version of the child behaviour profile: Turkish girls 7 - 12, *F. Akkök, & P. Askar* 731
Neurophysiological basis of cognitive abilities, *A.N.Lebedev, A.V.Markina, & I.V. Maltseva* 734
Guessing-proneness, behavioral and cognitive impulsivity: Sex differences, *D. Bratko, K. Squarcy, P. Zarevski, P. Kovacevic, & S. Logar-Djuric* 735
Evaluation of training programmes: The state of the art, *J. d. Koning* 739
The assessment of written composition, *G. Engelhard* 741
The construction of a wide range reading comprehension test for adults, *W. de Geus, O.W.H. Henneken, & P. Reitsma* 744
Psychometric analysis of the Turkish group assessment of logical thinking (GALT) test, *G. Berberoulu, & F.J. Oort* 748

466

Adjusting for measurement unreliability in multilevel models, *M. Yang, G. Woodhouse, H. Goldstein, H. Pan, & J. Rasbash* 751
Applications of multi-group IRT models, *M.F. Zimowski, R. Darrell Bock* 753
On the existence or otherwise of compositional effects in multilevel modelling of school effectiveness, *D. Hutchison, & I. Schagen* 756
An application of a multilevel model in the analysis of panel data, *M.W. Vooijs, C.M. Koolstra, L.J.Th. van der Kamp* 759

Posters
Problems in the evaluation of the assumptions of the model using box-plots, *A.M. Lopez, & M. Morales* 762
Uses of qualitative/qualitative terms in educational research, *M. Morales, & A. Moreno* 764
New means of feedback evaluation in research on education: Research methodology and evaluation research, *K.R.G. Ulbricht* 767

Symposia
Learner assessment and evaluation: New methods, *W. Wijnen, F. Dochy,& I. Pychova* 771
Problems in measuring the reliability of national curriculum assessment in England and Wales, *I. Schagen* 776
The Dutch national assessment program in education: Aims and scope, *J.M. Wijnstra* 782
Towards a student monitoring system in primary school, *J v. Bergen, L. v. den Bosch, T. Eggen, R. Engelen, P. Gillijns, F. Kamphuis, & F. Moelands* 786

Chapter 6
TEACHING AND TEACHER TRAINING 789

Papers
Teaching and teacher training, *J. Pieters, & B. v. Hout-Wolters* 791
The professional development of teachers in health education: issues in delivery of inservice provision in the United Kingdom, *G. Heathcote* 793
Development of an instrument for assessing classroom performance in a faculty of medicine, *Th.J. ten Cate, & J.A.L. van den Bijtel* 796
Implementing knowledge about language in the English national curriculum, *J. White* 800
Teacher problems and need for in-service training, *Th. Houtveen, & A.L. van der Vegt* 802
Assessment of starting competences of student-teachers, *R. d. Jong* 806
A study on the admission standards for teacher training programs, *F. Paykoç, & A. Ok* 809
Control and discipline: a problem for beginning teachers?, *W. Taylor, J. Draper, & H. Fraser* 813
Improving the teaching of beginning science teachers: a study of expert and novice classroom practices, *A. Cachapuz, & M. Jorge* 816

467

Introducing the reality shock into preservice teacher education, *C.P. Koetsier, J.S. t. Brinke, Th. Wubbels, & C. v. Driel* **820**
A typology of student teachers' coping reactions, *W. Admiraal, F. Korthagen, & T. Wubbels* **823**
Expert and novice history teaching: eight differences between experienced and beginning history teachers, *E. Ropo* **826**
The state and the teachers: problems in teacher supply, retention and morale in England and Wales, *G. Grace* **828**
The metacognitive teacher, *M. Cardelle-Elawar* **829**
Moving towards teaching: a model of teacher development, *R.G. Elliott* **834**
Teachers' evaluation of training and implementation of a new field study in the curriculum, *C. Bruinsma, & P. v.d. Jagt* **837**
Teaching and teacher training, *J. Lowyck* **840**
An evaluation of a national training program on educational leadership for primary schools, *C. Matthijsen, H.P. Brandsma, & W.G.R. Stoel* **844**
The identification of a common core curriculum in the training of teachers, social workers, health workers and community educators, *M. Hough* **846**
Access and transition into teacher education in an increasingly ethnic diverse society, *J.R. Meyer* **849**
Schoolleaders in training: an international comparison, *Q.H.J.M. van Ojen, & S. Karsten* **851**
Comparing modes of teacher education: insights into the process of a research project, *L. Barton, E. Barrett, G. Whitty, J. Furlong, & S. Miles* **854**
The 21st century Romanian teacher - a perspective viewpoint on his professional development and moral changes, *A. Dumitrescu* **856**
Initial formation of teachers at secondary educational level in Spain: an unsettled challenge, *V. Fernandez Benitez, & J. Mª Poulussen v. Balen* **858**
Counseling in secondary education after the implementation of the common curriculum, *C. Vermeulen, K. Bos, & L. Cremers* **862**
The evolution of work satisfaction, work load, burnout and feelings of a severe job crisis amongst Dutch teachers, *M-J. d. Jong* **865**
A methodology for eliciting different types of concept maps, *N. Siregar* **867**
A model of how students learn how to teach as the meaningful acquisition of professional activity, *G. Harvard, & R. Dunne* **871**
Developments in strategic reading instruction: effects of a training of teachers, *P. Reitsma, & S.E. Sliepen* **874**
The contribution of journal writing to student teachers' professional learning, *C. James, & J. Calderhead* **878**
The teaching profession: teachers' representations of their position in social space, *A. de Lima* **880**
The study of teachers' professional biographies: methodological evaluation, *G. Kelchtermans, & R. Vandenberghe* **882**
The professional development of primary school teachers, *M. Clement, & R. Vandenberghe* **885**
Career story, self and subjective theory: a biographical approach to professional development, *G. Kelchtermans, & R. Vandenberghe* **888**
Primary school teachers between autonomy and collegiality, *K. Staessens, & R. Vandenberghe* **891**

The training of university training staff: the Spanish challenge, *B. Antoli, F. Cerveró, & F. Pavia* 894
Prospective mathematics teachers' attitudes toward computer and computer assisted achievement in solving max-min problems in calculus, *Z. Toluk, & Y. Ersoy* 896
Evaluation of a pre-service programme based on direct instruction, *S. Veenman, Y. Leenders, P. Meyer, & M. Sanders* 900
The etobicoke classroom activities project: the meaning of educational change, *R. Volpe, R. Fleming, & A. Benins* 904
On the way towards active learning and developing worklife: development of teaching in Finnish businessmen's commercial college, *A. Kauppi, & S. Peisa* 905

Posters
Interaction styles and cognitive development, *M. Lera, & A. Aguilera* 908
Cognitive activating patterns: an instrument for the assessment of interactions in the classrooms, *J. Mora, & A. Aguilera* 911
Dynamic collaborative approaches to staff development review and needs identification at school and education authority level in Scotland, *A. Naylor, & J.P. O'Brien* 914
Creative climate in teachers' group: barriers and changes of innovations, *D. Kusá* 916

Symposia
Training for competence in teaching - assuring competent performance in practice, *F. Kieviet, & H.H. Tillema* 919
An overview of ten years of research in three Dutch research programmes on teaching and teacher education, *F. Korthagen, & Th. Wubbels* 930
Interpersonal teacher behaviour, *N. Verloop* 934
Perspectives on knowledge bases for teaching, *N. Bennett, C. Carré, & E. Dunne* 939

CHAPTER 4

EDUCATIONAL TECHNOLOGY

EDUCATIONAL TECHNOLOGY: INSTRUMENTATION AND IMPLEMENTATION

J. Moonen, The Netherlands, I. Stanchev, Bulgaria, & B. Collis, The Netherlands

Learning; teaching; the processes of inquiry, synthesis, and communication of concepts and understanding -- all of these areas are being significantly affected by the changing society in which we live. One of the characteristics of this society is the rapid growth of information and information technologies. The purpose of Stream 4 is not to discuss these new technologies in themselves -- many other specialist conferences have this purpose -- but instead to highlight significant research, predominantly European, relating to the nature of the learning experience in learning and social environments in which new technologies play a part. The research reported in Stream 4 covers a broad spectrum of perspectives relating to learning environments in which interactive, computer-related technology is a component. These perspectives range from empirical investigations of the implications of design decisions in electronic educational instrumentations for different groups of learners in different types of learning settings, to theoretical considerations of new paradigms for the organization of education and learning emerging in a so-called "Information Age". Particular attention is given to research methodologies appropriate to current learning situations involving computer-related technology, and to the new paradigms for teaching and learning which may be also emerging.

The keynote speech for Theme 4 is actually two interrelated speeches, presenting different perspectives on the roles of media and human actors when multimedia courseware is used in a variety of learning contexts. Prof. Robert Lewis and drs. Jan Schoenmaker amplify different approaches to the research process in this type of complex human-and-resource setting.

The remainder of Theme 4 is organized in symposia, groups of interrelated papers, a poster session and a closing session. The two symposia both focus on the challenge of implementing adequate and effective research methodologies for studying the impact of technology-enriched school environments over extended periods of time. The first of the symposia presents a comparison of methodologies for such research as have been conceptualized and implemented in a number of different European countries; the second symposium examines in more detail the methodologies and results from one such project, in The Netherlands.

The remainder of the papers in Theme 4 are grouped in 15 clusters, each cluster relating to an aspect of the overall Theme. Some of the clusters relate to research

focussed more specifically on new educational instrumentation itself, as in the subthemes relating to "Computer Simulations: Perspectives and Applications," "Research on Interactive Video," "New Developments in Electronic Learning Resources," and "Educational Software Evaluation." Other clusters are more focussed on research relating to the impact of new technologies in particular learning contexts, such as in relation to curriculum areas ("Applications of Information Technology in Science" and "Second Language Learning and Technology") or in different learning settings ("Technology and Classroom Processes," "Computer Implementation Perspectives", "Perspectives on Technology in the School"). Other clusters relate to more specific aspects of computer-related educational technology in learning and the organization and interpretation of information ("Strategies for Information Handling, Retrieval, and Organization", "Applications of Decision Support Systems"); to research relating to the educational implications of telecommunications ("Organizing Telecommunications for Learning"); and to overviews of research from different national and international perpsectives ("International Perspectives on Information Technology in Education.") Finally, research related to more general issues regarding the impact of technology in learning are the focus of the clusters "Perspectives on Effectiveness and Impact of Educational Media" and "Learning in a Technological Context."

Thus, Theme 4 deals with research from many perspectives -- methodological issues, conceptual frameworks, case studies of naturalistic settings, action research, instrumentation - focussed and implementation - focused -- but sharing as a common feature the disciplined examination of learning in technology-augmented settings.

CASE STUDIES OF CLASSROOM PROCESSES USING GEOGRAPHY SIMULATIONS

Deryn M. Watson, Centre for Educational Studies, King's College London, United Kingdom

Abstract

Longitudinal in-depth case studies of the use of IT in geography classrooms were undertaken as part of a large national project, 1989-1991. Research using ethnographic method explored the use of third/developing world simulations with 12-14 year olds. The simulations were found to be good motivators which heightened pupils interest, enjoyment and empathetical understanding and were associated with increased concentration and sustained effort. These findings support the role of enquiry focussed simulations in the stimulation of discourse and process based learning and thus the generation of a functional learning environment.

Introduction

The ImpacT study, an evaluation of the impact of IT on children's achievements in primary and secondary schools commisioned by the Department of Education and Science (UK) was conducted by researchers at King's College London in the period !st January 1989 - 31st December 1991 (Watson 1992b) The research focus was on pupils learning involving IT in the subject areas of mathematics, science, geography and English. As part of the project, longitudinal in-depth case studies were conducted; the author undertook the geography case studies and this paper reports on some of her research findings of classroom activities when using simulations.

Method

The researcher employed ethnographic method, using qualitative data gathered from observation and recording, formal and informal interviews and documentary evidence (Hammersley and Atkinson 1983). The specific aim was to get as close as possible to the classroom processes, where the teachers were known users of software. All classes involved were in co-educational comprehensives in different geographical and social areas.

Geography was one of the subjects chosen for study because it has a history of extensive software development and use over the last fifteen years (Watson 1992a). The whole geography case reported on a wide variety of issues related to

learning, teaching and organization/policy within the department and school. This paper focusses on the classroom processes when simulations were used.
Five sets of geography lessons, all with 12-14 years olds, were observed in four classes when four different simulations, covering third/developing world studies, were used. Although the context differed, the software displayed similar characteristics, that is they engaged the pupils in role playing activities in which they had to experience the life of Indian farmers, or African nomadic cattle herders, over several seasons or years. In each case the class was prepared beforehand with material relating to the topic. All teachers used the simulations as illustration of a response to seasonal weather patterns. Two of the classes moved to a computer room; the other two operated a cafeteria system using machines in the classroom. In all classes the pupils worked in groups, two were self-selected groups and two selected by the teacher.

Results

All four different classes observed responded to these simulations with a high level of *interest and involvement*. There were five main indicators of involvement.
Firstly, the general level of noise in the class rose and was tolerated by the teacher, compared with that allowed in their "normal" classroom. This level of noise was sometimes disturbed by a louder exclamation when a group saw the results of a particular decision. Secondly, the noise factor did not disturb the teachers because the pupils were so clearly talking on task; teachers tolerated this higher level of noise because they had expressly used the simulation in part to encourage group work and collaboration. Thus this discourse was welcomed. The only time that the researcher experienced the same level of noise and discourse in a geography class was when a group decision-making exercise was used, although it must be remembered that only two of these classes were observed over a substantial period of time.
Thirdly the involvement did not wane as the lesson progressed. In the classes in the computer room, pupils maintained the same level of involvement in the same activity throughout a one hour period. Again this did not usually occur in their geography classes, when the activities were usually broken up into different and clearly defined sections. Involvement was maintained even in the class where there were hardware problems. Fourthly, this involvement was serious. The pupils were not treating the lesson as a "doddle" in which they did not have to be involved. Interestingly in two groups from two different classes there were incidents in which members of the group chastised another member for not taking the work seriously and not therefore thinking carefully enough about their input.
Fifthly, pupils were often reluctant to finish the session even when the bell went and the next period was lunch or break. Some pupils sighed and requested a chance to continue on another occasion. There was a significant amount of regret from the two teachers who had whole classes in the computer room that booking problems prevented them being able to do this. The classes who used one or two machines in their own classroom, also had their time on the simulation reluctantly curtailed. Talk on task by the pupils occurred as they were leaving the room and going down the stairs. One class became very animated as they exchanged experiences about what had "happened to them". Animated talk on task was not the norm amongst pupils leaving their "ordinary" lessons.

A significant amount of *empathy* developed during these lessons. When asked what they had learnt, one group of pupils, John, Sam and Kit expressed themselves thus: "How hard it is. We didn't know you had to eat thorny scrub. We started panicking when they (the cows) all started dying. We started with 20 but went down to 8. (R: "Why did you panic?) "Well, we would all die if the cows died, it was really worrying." All three boys further explained the seriousness of the situation and how there was no "way out" if things just went on getting worse. This empathetic involvement was not just momentary but was reflected in the class writing that John did a week later on "the life of a nomad".

Pupils working on the Indian Farmer simulation discussed the problems of large numbers of births in an area where once a child is eight they can work fully in the field and productivity increases. The poverty trap and downwards spiral of health if the family was not able to harvest enough grain resulted in comments on the real differences in life for those farmers and the dilemmas they faced. Conversations during the simulation included "Oh God, not another baby" and "Sir, we've got ourselves out of debt". Comments afterwards included "It really made you think about the problems they face. Getting into debt...the reality of children dying...my wife died....had a new wife and then there were lots more children and so more food needed". Another said "I learnt quite a lot. I didn't realize there was a cycle in everything they do. I suppose our life has cycles, but its more like doing different things every day....like lessons". The direct involvement and association with the lives they were exploring was noticeable. When asked six months later what they remembered doing, they had a very strong recall of the simulation and could describe what "happened to them" in some detail. Thus involvement and vivid experiences were not transient.

There was a universal sense of *enjoyment* in the geography classes when they were using simulations. They provided a *challenge* because they taxed the pupils knowledge base by placing them in a context where the detail was greater and more complex than hitherto experienced. The emergence of the detailed implications of this complexity, such as with the complex health factor which influenced the overall progression of an Indian Farmer or the choice and combination of resources and activities for an African villager, suggested that this enquiry work was inherently challenging and involved more real choice for the learner. Many pupils said in effect "I learnt more than I thought there was to learn about this topic".

In one class, although using the simulation came at the end of a substantial period of studying the savanna cycle, the cycle of life of the nomads, and the interrelationship between the two, some pupils displayed an understanding of greater detail and "connectedness" of factors than before. Alan, on initial questionning about whether he had learnt anything different about the nomads when using the software, said "not a lot". But he then went on to say that he was surprised about the influence of the tsetse fly. Although he knew about them, "I didn't know that they could sort of hurt the cows". Graham with whom he had worked, said he could "see now why they should go in to town as if they sold butter it meant they could buy some more cows". Alan then said that it was a problem because "every time you moved to grass by the time you got there it was scrub because it got too hot". Pindar "didn't realise it rained so much in the summer, and you do get a lot (of grass) when it rains. But we didn't eat much of this 'cos it was in

the wrong place". Penny didn't realize "that the food was so bad for the animals; if they stopped to graze and the grass stuff was not good then, two cows died". Nicky and Lorraine who were working with her were surprised at how many kilometers (427) they travelled in the year. The combined effects of the impact of the seasons and travel had clearly not impinged upon them in the same way from their earlier class based exercises. The level of supporting detail that was added to their more generalised knowledge was impressive and was reflected in their subsequent writing. This is a reflection of the richness of the overall experience and relates to the issues of purpose driven exploration.

Much of the content of the normal geography lessons, not using IT, was based on presented material, and the choices the pupils had to make tended towards how much detail, out of that already presented, to include in their own work. The choices in these simulations were different. Pupils were having to sift through the detail for coherence as well as relevance. The overall impression was that the complexity was attractive and real to the learners. Whereas in normal lessons the material was presented and discussed at the beginning of the lesson, thereby in a way setting the agenda and framework for their subsequent writing, the lessons using IT were more truly *exploratory*, whether of ideas or presentation. And this exploration was in the hands of the pupils, though guided by timely teacher intervention.

Discussion

Although this case study work is limited in scope, some clear findings emerged that illuminate how teachers and pupils actually use computers in their classroom. Thus, computer simulations were found to be good motivators which heightened pupils' interest and enjoyment, and enabled pupils to develop a high degree of empathy with the topic under study. This motivation was intricately linked with increased concentration and sustained effort. By giving pupils the opportunity to work in an open ended way, simulations enabled pupils to become involved in more complex and challenging learning situations than normal.

Breuer (1985) relates the attributes of simulations to the concepts of cognitive complexity and conceptual mapping through the opportunity of choice simulations provide to the individual student to search, enquire, hypothesize and try out, and to work as a member of a team. In contrast Chandler (1992) questions the value of simulations as he claims they are removed from the existential context of the human experience they are intended to represent. He suggests that simulations mechanize experience; indeed he asserts that simulations are asocial, ahistorical and amoral. This geography study however provides no support for the contention that pupils themselves are engaged in a false or mechanistic activity. Rather these classrooms relate more closely to the functional learning environment of Newman (1990) that is predicated on the assumption that children are intrinsically motivated to work on tasks that are meaningful to them; the use of these third world simulations supports his statement that simulations are important tools for the solution of an interesting variety of real world problems.

Edwards and Mercer (1987) argue that discourse, rather than being just a window upon children's thought processes, is itself an educational reality. Teachers and children engage in constructing through context and continuity, a shared account, a common interpretative framework for curriculum knowledge. The motivation and

interest reported here, which reflected a more challenging working environment, supports such analysis where the nature of the software and the curriculum agenda set by the teacher places an emphasis on the decision making and discourse during the interaction.

References

Breuer, K. (1985). Computer simulations and cognitive development. In K. Duncan and D. Harris (Eds) Computers in Education North Holland: Elsevier Press.

Chandler, D. (1992). The purpose of the Computer in the Classroom. In J. Beynon and H. Mackay (Eds.) Technological Literacy and the Curriculum. London: Falmer Press.

Edwards, D., & Mercer, N.M. (1987). Common Knowledge: The development of understanding in the classroom. London: Methuen.

Hammersley, M., & Atkinson,P. (1983). Ethnography: Principles in Practice. London: Tavistock Publications.

Newman, D. (1987). Functional Environments for Computers in Education. In R.D.Pea , & K. Sheingold (Eds.) Mirrors of the Mind: Patterns of Experience in Educational Computing. New Jersey: Ablex Publishing Corporation.

Watson , D. M. (1992a). The Computer in the Social Science Curricula. International Journal of Educational Research, 17, No. 1.

Watson, D.M. (Ed.) (1992b). *ImpacT: an evaluation of the impact of IT on children's achievements*. London:Department of Education and Science.

HOW COMPUTERS ARE USED IN INSTRUCTIONAL PRACTICE: FOUR CASE–STUDIES IN A DUTCH SECONDARY SCHOOL

Wim Veen, Fred Vogelzang, Irma v.d. Neut, & Peter Spoon, Institute of Education, University of Utrecht, Utrecht, The Netherlands

Introduction

The Dutch government has been encouraging the use of computers in secondary education since 1982. Hardware has been distributed free to all schools, software has been developed and made available at very low prices, and money is being allocated to schools on an annual basis. The ultimate aim of the government's efforts was that the computer should be used in up to 5% of all secondary school lessons. It has become clear that the results are disappointing. Only 9% of the teachers in secondary education use computers from time to time (Min. of Educ. 1990). The question arises as to why the national program has had such limited results.

International research has shown that the teacher plays a central role in educational innovations (Chandra 1987, Fullan 1982, Olson 1988, Wiske 1988). Despite the national in–service training program for teachers, it must be said that little attention has been paid to the teachers' role. The program was mainly focused on hardware and software. Thus national programs hardly took into account what teachers usually do in their classrooms. Teachers were not provided with computers in those places where they usually work. A teacher–centered approach is most common and teachers usually prepare lessons at home. The use of computers forced teachers to depart from their routines and simultaneously to change several aspects of their teaching.

The authors decided to investigate the teacher's role in the use of computers in their instructional practice. Therefore it was necessary to provide them with hardware. Through a joint study agreement with IBM we provided four carefully selected teachers with a computer at home, one computer and a transviewer in their classroom and a computer lab with thirteen computers. The teachers did not have other facilities during the research study.

The main question to be answered was: 'What are teachers going to do with the computers at school and at home and can their activities be explained?'

The research

The aim of this research study is to provide an understanding of how teachers actually use computers in their teaching and why they are doing so. In order to get an overall picture of what influences are at work in the uptake of computers, an in–depth description of the instructional practice of teachers was considered necessary. For this reason a case–study was used.

We used different sources of data. Teachers kept diaries of every computer (–related) activity. The diaries were semi–structured and teachers sent them to the researchers every week. In connection with the diaries, teachers were interviewed at school every fortnight. We observed computer assisted lessons in the classroom and in the computer lab. Finally we used in–depth interviews every six months with the teachers, the principal, the technical assistant and the IT–coordinator.

In order to eliminate bias in our data, we each read the interview transcripts, and rotated in our classroom observations and interviews with teachers. We discussed results frequently to reach a high degree of inter–subjectivity.

From the fall of 1989 until January 1991 all computer (–related) activities were carefully described.

The school and the teachers

The four teachers in French, English, Geography and History all worked at a medium size secondary school (800 students) in an urbanized region in the centre of the Netherlands. They were teachers who had a good reputation and had about 20 years experience except the history teacher who had 5 years experience. Roughly 80% of their lessons were teacher–centered. Compared to 95 other Dutch teachers on eleven characteristics (Brekelmans 1989) they appeared to be rather average teachers, although the English teacher scored very low on willingness to innovate. Certainly they did not belong among the computer pioneers and did not have experience with computer assisted lessons. However, two of them used a computer at home for word processing.

Their views about the use of computers in teaching differed. The Geography and French teachers had positive views. They expected to use a lot of computers for their teaching. The History teacher can be characterised as critical but positive (Chandra 1987). He had some critical comments about the way they should be used. The English teacher had several worries and fears (Chandra 1987) about the use and implications of computers for his teaching situation.

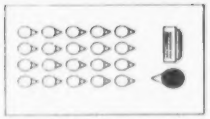
The results

After a year and a half of collecting data we had a clear view of how teachers used their computers at home and at school. Two research questions could be answered. The questions were defined as follows:

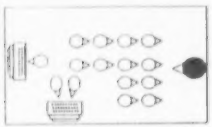
- Which activities do teachers undertake at school?
- Which activities do teachers undertake at home?

We found a large variety of computer assisted lessons (CALs) in observing the classroom activities of the teachers. Thus we considered it necessary to categorise the different settings. Categories were based on the configuration used and the pedagogical setting. In this way we distinguish seven different types of CALs, as shown in figure 1.

Fig. 1

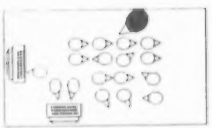


Cal 1: the electronic blackboard
The traditional teacher-centered classroom situation. The computer is used in combination with a transvie wer for instructional purposes.



Cal 2: working apart together
Some students are working indepently in the teacher's classroom on special tasks. Tasks can be initiated by teacher or requested for by the students and may last for just a part of the lesson. The teacher is working with the others in an instructive way.

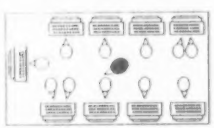
Cal 3: rotating groupworking
All students work together on a common task. Task may be in connection to the computer or not. they rotate in using the computer. The teacher supervises by walking around.



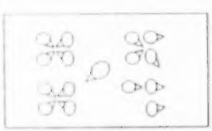
Cal 4: artificial classroom reduction
All students work individually or in groups. Besides the computer students may use other media, like audiotapes or textbooks. The teacher may work with individuals or may train certain skills with a small group.

Cal 5: volunteer working
Students ask for special tasks (training exams, vocabulary etc.) and work with the computers in recreation time or after schooltime. The teacher is sometimes present for help.

Cal 6: computerlab without support
All students work on the computers individually or two by two. The teacher supervises by walking around answering questions etc.



Cal 7: computerlab with technical support
Classes are divided in two parts. One part is working in the computerlab supervised by the technical assistant, the other part of the class is working with the teacher in his classroom/ sometimes the teacher jumps in the computerlab to give some help.



As far as their computer activities at school are concerned, the teachers appeared to be quite different. As shown in figure 2, not only did the number of computer assisted lessons the teachers organised vary a lot, the types of computer assisted lessons differed too.

Fig. 2
Computer assisted lessons of the four teachers during the research study.

Types of CAL's	French	English	Geogr.	History
CAL 1	12	9	28	14
CAL 2	19	5	-	-
CAL 3	5	-	-	-
CAL 4	30	-	4	-
CAL 5	17	-	14	4
CAL 6	23	2	13	15
CAL 7	19	2	6	-
Total	125	18	65	33

The French teacher used computers most of all and he used them in all the different types of CALs. He was running, as he called it, 'experiments' with these different types of CALs. Although he used computers very frequently, it must be noticed that he mainly used them for three purposes: to give information to the whole classroom (CAL 1), for remedial purposes and for drilling grammar and vocabulary. Small groups or individual students worked on the machines during the lessons or in recreation time. They worked in the classroom as well as in the computerlab. This teacher thinks computers motivate students and lead them to work independently, which in his view should be the way for the future.
The English teacher on the contrary used the computers only in eighteen lessons. He stopped computer activities after some disappointing experiments during the end of 1989. He thought computers disturbed his classroom order too much and he attributed his negative experience to a lack of motivation on the part of the students and to the size of his classes. Moreover all the software he looked at did not suit the textbook he was using closely enough and, in his opinion, did not add any educational value to his lessons. Finally, using computers did not suit his teaching style. He considered computers only to be useful in very small classes of very motivated students.
The Geography teacher made use of his classroom computer and transviewer frequently. Moreover, after getting acquainted with a geographical database, he used this configuration often for just five or ten minutes to show some graphics. In that way he fitted the computer seamlessly into his expository teaching style. Besides this, he regularly went to the computer lab with his classes to let his students work on computer simulations.
he History teacher became aware of some advantages of computers in his lessons: computers stimulate an explorative learning approach and motivate students in history. Nevertheless he usedthe computer less as part of his regular routine than the French and Geography teacher. The first reason was a serious

lack of software in his domain. Secondly, although he saw the advantages of computers, he felt a preference for expository learning. He thought a history teacher should always explain the general socio–economic and political backgrounds to historical events. And this is what, in his opinion, was lacking in historical course–ware.

As far as the computers at home are concerned, all teachers used the computers frequently and intensively for word–processing. See figure 3. The four teachers considered this computer application as very useful and time–saving. This use appeals to the need for efficiency. In the other categories, differences in activities appeared.

Fig. 3
Computer activities at home. Categories have been determined in collaboration with the teachers.

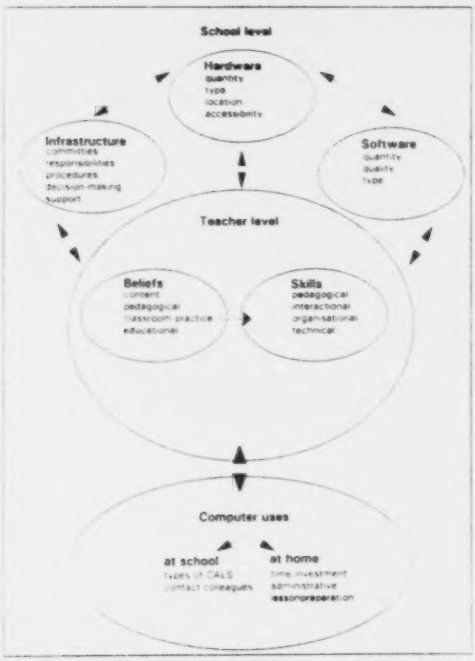
Activities	French	English	Geogr.	History
Brousing courseware	7%	2%	19%	20%
Filling courseware	10%	19%	0%	0%
Making pupils' materials	23%	61%	41%	36%
Administrative activities	2%	25%	2%	15%
Related activities	39%	3%	17%	24%
Lesson preparation	19%	0%	22%	5%
Total in %	100%	100%	100%	100%
Total in minutes weekly	139	226	143	153

Both language teachers filled in course–ware with excersises and sentences they liked to use. In Geography and History this kind of course–ware was not in use. They spent more time in browsing through new software to get acquainted with it, and use it in the classroom possibly. Administrative uses were experimental and depended largely on individual preference. As far as the school was concerned the use of administrative course–ware was not compulsory. The preparation of lessons took the Geography and French teacher about one fifth of the time spent at home. The Geography teacher prepared his lessons carefully, specially in the beginning, as he did not want to trust to luck. On the contrary, the History teacher preferred not to spend too long in preparing his lessons and liked to try out things and see the results in practice.
Although the teachers showed significant differences in computer uses in the classroom, they all spent about two and a half hours a week at their computers at home. The teachers felt this to be an acceptable time in their normal weekly work load.

Towards a descriptive model
Now that we know how much time teachers spent working on the computer at home and the intensity and variety of CALs at school, the question remains how the uptake of computers by the teachers can be explained?

To answer this question we searched in all interviews and diaries for the arguments and reasons teachers had used. In an intersubjective approach, those arguments and reasons were confronted with the views of the researchers. This confrontation has led to a number of related arguments and reasons, which we call here 'factors'. It had become clear from the original statements of the teachers that the factors appeared not to be isolated from each other: on the contrary they were strongly interrelated. Therefore we developed a descriptive model of the factors that were at work in the introduction of the computer by these four teachers.

Fig. 4
Descriptive model of factors affecting computer uses of the case-study teachers.



We discerned two major categories of factors: factors at school level and factors at teacher level (Fullan 1982).
School factors played an important role. Firstly, hard– and software had to be easily accessible and secondly, the computer at the teacher's home appeared to be essential. As teachers said, it helped them doing what they are used to do: preparing their lessons at home. Furthermore the principal played an important role. He provided the necessary technical, financial and, last but not least, moral support.

However, teacher factors outweighed the school factors. These were divided in two categories: beliefs and skills (Olson 1988). Most important of these were teachers' beliefs regarding to what should be in the curriculum (content) and the way in which the subject should be taught (pedagogy). Besides, teachers had beliefs regarding their role in the classroom and the corresponding classroom activities. Finally they had personal views on education (educational) and on themselves as teachers. The skills of teachers were related to their competence in managing classroomactivities, to pedagogical skills and, less important, to the technical skills. They can briefly be described of as 'routines'. In our research, beliefs seemed to determine or influence skills more than the other way around. The English teacher believed that teacher–centered classroom situations lead to more effective learning, so he was not willing to develop skills for pupil–centered classroom situations. On the other hand, the French teacher was convinced of the advantages of group–work, without having the teaching skills, so he was eager to develop the necessary skills.

Conclusions and recommendations

Results of case studies cannot be generalised. Thus the conclusions below must be considered in the context of this case study. One of the most important conclusions of the research was that the uptake of computers in the classroom was strongly affected by the beliefs and skills or 'routines' of teachers. When the use of computers and software made the teachers deviate too much from their teaching ideas and routines, they decided not to use them at all. Their beliefs in regard to the content and the pedagogical strategy of the subject were the strongest factors. Technical skills were far less important.

The researchers recommend that the first introduction of computers should fit as much as possible into the day–to–day 'routines' of teachers. Using computers should fit into existing skills of teachers and once they have gained experience after two or three years, teachers can gradually enhance their routines and handle more complex applications of information technology. In this way teachers' skills will begin to change their beliefs. It is our belief that more than 10% of the Dutch teachers want to use information technology but they do not want to be forced to turn their teaching routines upside down. The researchers think that in that respect teachers do not differ from any other group of professionals.

References

Brekelmans, J.M.G., (1989). Interpersoonlijk gedrag van docenten in de klas, p. 157–161, PhD diss. Utrecht.

Chandra, P., (1987). The implementation of computers in a secondary school: A case study of teachers' perceptions about computers in teaching within the social organisation of a comprehensive school, diss. King's College, London.

Ministry of Education (1990). Evaluatierapport computers in het onderwijs, Den Haag.

Fullan, M.G., (1982). The meaning of educational change.

Olson J., (1988). Schoolworlds – Microworlds: computers and the culture of the classroom, Pergamon.

Veen, W., et al, (1991). De realiteit van computergebruik in de lespraktijk, vier case studies op een havo/vwo school, de Lier.

Wiske, M.S., et al.(1988). How technology affects teaching. Educational Technology Center Harvard Graduate School of Education, Cambridge, Mass. en Education Development Center, Inc. Mass.

PARTICIPANT BASED FORMATIVE EVALUATION OF A COLLABORATIVE TECHNOLOGY PROJECT: WORKING WITH IBM AND THE UNITED STATES DEPARTMENT OF EDUCATION

James E. O'Connor, California State University, USA

Evaluating technology programs is an arduous and intensive task. Especially important is formative evaluation during the initial or start-up year for multi-year projects. Participant-based methodology is an evaluation model that can be employed to provide effective formative evaluation to project leadership and project participants enabling them to adequately accomplish project objectives. This paper provides insight into participant-based formative evaluation by sharing the strategies, techniques and results from the first two years of the evaluation of the Mathematics and Science Partnership Project (MSPP) funded by the United States Department of Education. This project operates in conjunction with the Multimedia Mathematics and Science Education Specialist Project funded by the IBM California Education Partnership (ICEP).

Based in Bakersfield, California, the MSPP is a cooperative agreement between California State University, Bakersfield (CSUB) the Kern High School District (KHSD) and ICEP. It is designed to educate and empower high school teachers in order to raise expectations and facilitate improvement for mathematics and science instruction.

MSPP teachers use networked IBM computers as well as one multimedia teacher workstation within each of their classrooms. Teachers within the project are charged to develop multimedia lessons, materials, and activities for use in integrated mathematics and science programs.

The overall approach to the evaluation of the project is participant-based. This approach allows for the flexibility necessary for evaluating such a complex project which involves six different school sites and thirteen mathematics and science teachers teaching a variety of subjects. The various activities and strategies employed within the project evaluation during the first two years of the project are highlighted in this paper/presentation. The results, as detailed in the formal yearly formative evaluation reports, from both the first and second year of the project are reviewed and discussed.

Objectives

The objectives for this paper/presentation are to inform the audience about the various strategies and techniques used within a participant-based formative evaluation, specifically those used during the first two years of the MSPP. This

includes a detailed discussion of how the evaluation actually progressed and how participant involvement was established and utilized to structure the evaluation process.
Another important objective is to demonstrate how both qualitative and quantitative data were aggregated, synthesized and analyzed in order to enhance the effectiveness of the project. This ongoing process provided project leadership with formative evaluation with the goal being to empower the participant teachers as they pursued the project objectives which focused upon the development of new and innovative curricula.
Finally, another presentation objective is to highlight the suggestions that were made to the project leadership at the end of both the first and second year of the project via the end-of-the-year formative evaluation reports. These suggestions were constructed from an intensive participant-centered process in order to insure the future success of the project.

Perspective
The overall perspective of this paper/presentation is to discuss evaluation from a formative viewpoint. More specifically, the perspective emphasizes a client-centered and participant-based approach to evaluation. Furthermore, the techniques for the aggregation and synthesis of both qu~ and quantitative data is an important aspect of this presentation which has important theoretical considerations.

Methods, Techniques, and Data Sources
The techniques employed within this evaluation are many and varied. Initially the focus of the evaluation was on process as meetings were held to obtain feedback from the project participants including project leadership, school site administrators and teacher participants. This feedback was then aggregated and analyzed and subsequently again presented to the project participants for their reactions and suggestions. As a result of this process,the various strategies and methods used during the evaluation were identified.
Following these initial group meetings, individual conferences were held with each of the project participants including project leaders, project consultants, school site administrators and project teachers. One-by-one these participants provided feedback about the evaluation process and provided focus for specific methods and strategies that they would individually use within the evaluation process.
Some of the evaluation methods that were employed during the first two years included : teacher journals; teacher interviews; student interviews; teacher surveys; student surveys; classroom observations; examination and analysis of curricula, i.e.. lessons, that were developed; interviews with school site administrators; field notes; computer logs documenting use of the computers and computer networks by both teachers and students; comprehensive end-of-the-year questionnaires for project leadership, school site administrators, project consultants and project teachers; analysis of attendance data; pre- and post-tests comparisons of student achievement; and quasi-experimental teacher action research comparing students immersed in the technology enriched environments with students in control groups.

Results

The results of the first two years of interim evaluation focus around the formative aspects of the evaluation process. Once the data was collected, aggregated, synthesized and analyzed, the results were contrasted with the specific objectives of the project. This process demonstrated that several of the important project objectives had been successfully attained during this first two years of the project despite significant start-up problems at several school sites. Some of the successful accomplishments included the development of some highly innovative lessons using such tools as IBM Linkway and Math Toolkit. These successes also manifested themselves through increased student and teacher motivation and , in some cases, increased student achievement.

Additional findings were that the teachers were teaching in a much different manner. Their approach to teaching was becoming more student-centered as the teachers became more facilitators of learning rather than mere disseminators of information. Students frequently worked in cooperative learning groups where they could collaborate their efforts. This resulted in increased student problem-solving and higher level thinking abilities.

However, the evaluation process also demonstrated that several of the objectives had not been successfully met during the initial start-up year. The successful formative evaluation provided the necessary suggestions to enable the project to accomplish its objectives during the second yearof the project. As mentioned previously, the suggestions provided to the project leadership, resulted from the aggregation and synthesis of both qualitative and quantitative data via a highly participatory process.

Educational Importance of the Study

This study provides an in-depth examination of how to successfully employ participant-centered formative evaluation for highly complex technology-based multi-year projects. Further significance is given to this study in that the project involves collaboration among several partners including business,a government agency, higher education, and a public school system.

This study also takes on increased importance because it is based within a school system and geographical location which contains a highly ethnically diverse population. Much of the population has limited English language proficiency. A high percentage of the students and families of the students come from lower socio-economic backgrounds and many are new immigrants to this country. These variables added to the complexity of successfully implementing the school-based changes embedded within the project objectives and goals.

Finally, this study is important because it centers around the development of highly innovative integrated mathematics and science curricula which is actually being developed by classroom teachers who are empowered through the use of state-of-theart technology.

A COMPUTER BASED INSTRUCTIONAL STRATEGY FOR RESTRUCTURING STUDENTS' CONCEPTS OF FORCE

Paul Gorsky, Director National Museum of Science, & Menahem Finegold, Department of Education in Technology and Science, Technion, Israel

Abstract

This paper reports on the development of a prototype intelligent tutoring system consisting of a series of computer programs which simulate the outcomes of students' perceptions regarding the forces acting on objects at rest and in motion. A student is asked to set up a force array showing the forces he or she claims are acting on an object and is presented with a simulation of the resulting behavior of the system. This behavior may or may not be compatible with familiar behavior in the real world. An arrow-based vector language enables students to express their conceptual understanding and makes simple dialogue possible.

Introduction

A very large body of research has demonstrated that students' pre-instructional knowledge about force and motion is often logically opposed to the laws of Newtonian mechanics and is also highly resistant to change. In attempting to change students' intuitive beliefs about physical phenomena, many researchers have developed instructional strategies based on the generation of cognitive dissonance. One such strategy, outlined by Nussbaum and Novick (1982), includes three stages:

- an "exposing event" requiring student interpretation based upon preconceptions;
- a "discrepant event" creating conflict between exposed preconceptions and observed phenomena which cannot be explained;
- a "learning support system" to aid the students' ongoing search for a viable explanation.

This strategy, implemented in nine computer simulations portraying objects at rest and objects undergoing linear, projectile and periodic motion, included the following elements:

- a simple, arrow-based, student/computer language;
- a cognitive model to represent students' conceptual frameworks;
- a tutoring model to help students learn the new concepts.

Instructional Problems and Methodology

The central problem is to restructure students' scientifically inappropriate beliefs about force. To do so, we developed nine simulations which invite the student to state his or her per-ceptions of the force or forces acting in each situation. The simulations portray:

- A book at rest on a table;
- A pendulum bob at rest suspended from a string;
- A magnet at rest on a vertical iron surface;
- A book sliding on a frictionless table after being struck by a spring;
- A book sliding on a table (with friction) after being struck by a spring;
- A pendulum bob experiencing periodic motion after being struck by a hammer;
- A magnet undergoing vertical linear translational motion on an iron surface after being struck by a hammer;
- A ball propelled straight upward after being struck by a piston ;
- A ball momentarily at rest at the height of its trajectory after being struck by a piston and propelled straight upward.

Students' perceptions about the forces acting on each of the objects are applied to the system and the physical outcome is simulated. The resulting simulated behavior may or may not be compatible with familiar behavior in the real world.

Exposing Events

For each of the simulations in turn, an "exposing event" is created. This includes a graphic and/or animated portrayal, a textual description of the system, and a question designed to elicit the student's understanding of the situation.

Discrepant Events

A "discrepant event" should clearly and dramatically negate the validity of students' misconceptions. Our programs show students the outcomes of their beliefs instead of simply informing them that their beliefs are incorrect. For example, if a student claims that no forces act on a pendulum bob at rest, then the string is severed and the bob is shown to remain in place, instead of falling.

The Learning Support System

The goal of the learning support system is to help students in searching for solutions to the anomalies presented. To this end, the system creates a profile of a student's explanatory frame-works, based upon performance within the program, and compares these profiles with that of an expert. Information needed to solve the problem is provided only if requested by the learner.

Research Design

Nine high school students in grades 9-12 were observed as they used the simulations to learn about force. Seven students had not studied physics and two had recently completed a course in mechanics. Typical of naturalistic inquiries, our observations and notes constitute a very large and amorphous data set. Findings are presented below from several different perspectives.

Research Findings

Creating dissatisfaction with currently held views

- Exposing events, using a simple arrow-based language, were effective in eliciting students' beliefs about the forces acting on objects at rest and in motion;
- Animated discrepant events enabling students to directly experience outcomes of non-scientific beliefs were effective in refuting these beliefs. Students experienced varying degrees of cognitive dissonance which led to schemata restructuring;
- Non-animated discrepant events (e.g. text messages) were not effective in creating cognitive dissonance.

Establishing initial intelligibility and plausibility of new scientific concepts

- The learning support system was generally successful in helping students perceive the intelligibility of the new concepts. In its present form however, the system is unable to help students perceive the plausibility of the new concepts;
- Researcher/student discussions, in the form of Socratic like dialogues, helped students perceive the plausibility of the new concepts;
- Although it is possible to develop computer programs to establish concept plausibility, we recommend that this task be assigned to instructors.

Student responses to cognitive dissonance

The degree of disequilibrium experienced by each student on viewing anomalies seemed to be a function of the prior knowledge held by the student.

- Anomaly, explainable either in terms of a current schema or as a result of making minor changes in a current schema, led to mild disequilibrium or none at all;
- Anomaly, not generally explainable in terms of a current schema, led to extreme disequilibrium.

Schemata transformations

- In learning about force, schemata transformation seemed to result from the interaction between what a student is taught and what he or she already knows;
- The major restructuring of central organizing concepts seems to be a gradual and reiterative process in contrast to the rapid change predicted by the catastrophe theory of learning (Gilbert and Watts, 1983);
- "Knowledge fragments" (Reif, 1987) apparently result from the inability of the student to make broad generalizations. Newly learned rules are apparently applied to the surface features of a system and not to the underlying processes.

Conclusions - Implications for Instruction

In this study, we have assumed that concepts such as force cannot be taught by the mere presentation of a logically organized body of knowledge, but must take into account learners' existing beliefs. The simulations do what a teacher cannot realistically do, namely, deal with each student and his/her explanatory frameworks on an individual basis.

References

Gilbert, J., & Watts, D. (1983). Misconceptions and alternative conceptions: Changing perspectives in science education. Studies in Science Education, 10, 61-98.

Nussbaum, J., & Novick, S. (1982). Alternative frameworks, conceptual conflict, and accomodation: Toward a principled teaching strategy. Instructional Science, 11, 183-200.

Reif, F. (1987). Instructional design, cognition and technology: Applications to the teaching of scientific concepts. Journal of Research in Science Teaching, 24, 309-324.

USING THE COMPUTER AS A HELP TOOL DURING LEARNING BY DOING

Marcel V.J. Veenman, Jan J. Elshout, & Janet G. van Hell,
Department of Psychology, University of Amsterdam,
Amsterdam, The Netherlands

The objective of this study was to determine whether students are capable of using computerized help-facilities during learning to solve thermodynamics problems. Students could voluntarily consult a help-facility furnished with different levels of help that incorporated a systematic problem approach. Thinking-aloud protocols were analysed on the quality of help usage and on the effectiveness of help requested as indicated by adopting the systematic problem approach. These measures of help usage were related to learning outcomes and student characteristics (intelligence and level of expertise). The results of this experiment may delineate some preconditions for the effectiveness of help tools in learning by doing.

Learning by doing in a semantically rich domain like physics requires an adequate level of domain specific knowledge and an effective problem approach (Jansweijer, 1988; de Jong & Ferguson-Hessler, 1984). Especially novices and less bright students are apt to strategic flaws (Veenman & Elshout, 1991). Lack of prior knowledge, both declarative and procedural, and lack of understanding often lead to a problem solving process characterized by a succession of impasses and subsequent repairs to those impasses (Brown & VanLehn, 1980; Jansweijer, 1988). This accumulation of impasse-repair sequences results in a disorderly and meandering pattern of problem solving activities, which constitutes an unfavorable context for learning. The central issue of this study is to explore whether students are capable of using the computer as a help tool for remedying their impasses and for adopting a Systematic Approach to Problem solving (SAP; see Mettes, Pilot, & Roossink, 1981).

The first objective of this study is to explore the amount and type of help requested spontaneously by students. Furthermore, we will examine the quality of the help requested (right type of help at the right moment) and whether help is used effectively. It is predicted that both high intelligent students and the more advanced students will utilize the help facilities more advantageously. The second objective of this study is to determine whether the received help leads to an improvement of the student's problem approach, as reflected in an accommodation of the SAP (Mettes et al., 1981). Finally, both the quality of help and SAP scores are expected to be related positively to learning outcomes.

Method

A CAI program was built that presented students with 20 problems in the domain of thermodynamics. Students had to solve these problems subsequently through the application of either the Boyle-Gay Lussac or the Poisson law. They were free to use the help-facilities of the program, accessible by a pulldown menu. The help-facilities were organized according to a step-by-step action plan that was based on SAP (Mettes et al., 1981). Any help given followed a sequence of subsequent steps that were concerned with: 1) reading the problem statement, 2) making a drawing, 3) constructing a scheme, 4) estimating results, 5) selecting a formula and converting units, 6) formula conversion, 7) performing standard operations, and 8) evaluation of the result. Additionally, four different types of help corresponding to increasing levels of support, could be requested. At the lowest level ("Clue") the students could ask for a clue concerning one specific step in the problem solving process. At the second level ("One step") the program could be requested to work out one of these specific steps completely. The third level ("Student performed") took the student by the hand during the entire problem solving process. At this level the student was instructed which steps to perform subsequently, however the activities for each step were to be executed by the student before standard feedback was given. At the highest level ("Computer demonstrated") all steps were comprehensively demonstrated by the program. Fifteen first-year psychology students, selected on being either high or low intelligent and on being either a novice or a more advanced student in physics, participated in the experiment. At the beginning of the experiment an introductory text about thermodynamics was studied by all subjects for 20 minutes. Next, a 40-item pretest questionnaire, tapping qualitative knowledge about thermodynamics, was administered. Working through the CAI program took about 3 hours, during which period of time thinking-aloud protocols were tape-recorded. Program traces with data concerning help requests, answers and time indications, were backed up for all subjects. After completion of the program a reshuffled version of the questionnaire was presented as a posttest. Also a series of 7 quantitative problems (equivalent to the ones presented in the CAI program) was given. After a two weeks delay both the questionnaire and another series of 7 problems were presented as retention tests.

The transcribed thinking-aloud protocols were analysed by two "blind" judges. For each subject the 20 problems were scored separately on measures for following SAP and on the quality of help requested (type, moment and utilization). Analyses were performed by means of Anova's (with intelligence and expertise level as independent factors) and correlational analyses.

Results and Discussion

Analyses of the computer traces on the amount and type of help requested, revealed that only 25% of the help requests were related to orienting activities (step 1-3), whereas 70% of the help requests concerned formula operations (step 4-7) and 5% concerned evaluation activities (step 8). More detailed analyses indicated that especially novices and low intelligent students requested more help in executing formula operations (step 4-7). Furthermore, students showed a significantly stronger preference for passive levels of help (One step and Computer demonstrated) instead of help levels demanding more effort (Clue and

Student performed). The students did not differ in the total amount of help requested. However, the scores for the quality of help requested that were judged from the protocols, clearly showed that the advanced students were more competent in using the help facility. Low intelligent novices scored rather low on quality of help requested.
The measures for following SAP, also obtained through protocol analysis, showed a significant increase over the 20 problems. Evidently, most of the students adopted a better problem approach while working through the program. However, high intelligent students revealed a stronger progression in SAP scores compared to low intelligent ones, regardless of the expertise level. An average SAP score that was calculated for each student, correlated .90 with scores for the quality of help requested.
The qualitative knowledge of students increased significantly between pretest and posttest. High intelligent students performed significantly better on all learning measures than low intelligent students, while advanced students performed better than novices. Moreover, high intelligent novices performed on a level equal to low intelligent, advanced students. Both the scores for the quality of help requested and the average SAP scores correlated positively with learning measures, whereas the amount of help requested correlated negatively with learning outcomes.
The results of this experiment warrant the conclusion that quality of help requested is a far more important predictor of learning than quantity of help requested. They also indicate that quality of help requested and SAP scores are strongly correlated and that the two of them are related positively to learning outcomes. Apparently, low intelligent novices in particular fall short of an adequate level of help utilization and, partly due to that, they also lack a systematical problem approach. The accumulation of knowledge deficiencies, limited intellectual resources and a poor problem approach is impedimental to the learning outcomes of low intelligent novices (Veenman & Elshout, 1991). Therefore, the results of this experiment may delineate some preconditions for the effectiveness of such help tools in learning by doing. Low intelligent novices may need more guidance in seeking adequate help at the right moment and clear-cut instructions of how to incorporate such a systematical problem approach. A recent study by Veenman, Elshout and Busato (in press) provided some evidence that instructions of a systematical problem approach may result in a higher level of learning performances.

References

Brown, J.S., & VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. Cognitive Science, 4 , 379-426.

Jansweijer, W. (1988). PDP: Een benadering vanuit de kunstmatige intelligentie van probleemoplossen en leren door doen in een semantisch rijk domein [PDP: An artificial intelligence approach to problem solving and learning by doing in a semantically rich domain]. Doctoral dissertation. Amsterdam: University of Amsterdam.

de Jong, T., & Ferguson-Hessler, M.G.M. (1984). Strategiegebruik bij het oplossen van problemen in een semantisch rijk domein: Electriciteit en magnetisme [Strategy use in problem solving in a semantically rich domain: Electricity and magnetism]. Tijdschrift voor Onderwijs Research, 1 , 3-15.

Mettes, C.T.C.W., Pilot, A., & Roossink, H.J. (1981). Linking factual and procedural knowledge in problem solving: A case study in a thermodynamics course. Instructional Science, 10 , 333-361.

Veenman, M.V.J. & Elshout, J.J. (1991). Intellectual ability and working method as predictors of novice learning. Learning and Instruction, 1 , 303-317.

Veenman, M.V.J., Elshout, J.J., & Busato, V.V. (in press). Metacognitive mediation in learning with computer-based simulations. Computers in Human Behavior.

EXPERIENCES IN PROFESSIONAL EDUCATION OF OPERATORS OF LARGE INDUSTRIAL OBJECTS USING KNOWLEDGE-BASED SYSTEM

Z.S. Hippe, Department of Computer Chemistry, Technical University, Rzeszow, Poland

Basing on the example of a large, fully automated chemical plant, the knowledge engineering program environment of a general use, SCANKEE, has been applied for training of operators (also called plant controllers) of large industrial objects, like production installations, power stations, coal mines, factories, mineral resources fields, etc. The main objectives of the developed system was to facilitate educational processes in a very complex high- tech environment and make them more effective. On the other hand, the proper use of the system promotes development of intelligence and learning in industrial organizations. The results of application of the system, gathered during education (training) of plant operators in the CS2-factory, are briefly dealt with.

Introduction

To begin with the problem of professional education of operators, let us to describe briefly some of the main difficulties of the automated production of CS2, based on carefully controlled burning of sulphur under high pressure, and in a very high temperature (exceeding several hundreds C the boiling point of CS2). In the case of malfunction of the pressured air installation, only 7 minutes are allotted to undertake and execute proper sequence of actions, whereas in the case of lack of cooling water, the operator has only 1 minute to break the production and exhaust the installation from the unused materials. The entire plant is permanently monitored on the control panel, watchfully observed by a plant controller on duty point, whose reaction in the case of warning signal issued by control automatics must be fast and correct. It should emphasized, that the CS2-production is performed in a very corrosive atmosphere, what may cause a malfunction of literally any of measuring devices on the installation (about 4000 of points where the automatics checks temperature, pressure, level, mass flow, or concentration, etc. [1]). For this reason, well educated plant operator should have profound factographical and procedural knowledge about the object being controlled, a very extended experience, and even the intuitive skill to recognize suspicious or unusual flow of some sub-processes despite of a fairly smooth and quiet performance of the control automatics.

Method

The combined method of computer-assisted education of industrial operators elaborated by us (using the own computer program system SCANKEE) has been directed towards finding answers to some important questions of CAI in its broad sense:

- how can knowledge be used intelligently, so it is applied properly in training and evaluating of operators of the very large high-tech industrial objects?
- what supports a sufficient, or an acceptable explanation for different classes (intellectually) of plant controllers?
- by what means the developed method may be combined with pedagogical and user knowledge, to implement intelligent information system for educational purposes on various levels?

Results

According to our approach, the training and learning procedure should be performed on three different levels. On the lowest level, which resembles "rote learning" (known learning procedure in artificial intelligence [2]), the user (trained operator) is confronted with the general scheme of the plant (layout, flow of information, etc.), having additionally the unique opportunity to enter into details of the object being regarded. Moreover, the user trained may trace all fine details of the installation, like: valves, measuring and controlling devices, and their current state. This level of training allows to check the actual understanding of the plant and its production; additionally, the operator becomes familiar with the general appearance and layout of the main control panel.

On the next, more advanced level of training, some questions are randomly generated by the system, pushing the user forward to the problem-solving and decision making areas. Usually, first are generated questions connected with finding causes of various (again randomly selected) malfunctions of the plant (installation). in a case of the correct location of a fault and its reasons, the last level of training procedure is executed. Now, the system generates a series of questions that leads to establish the proper response of the plant controller in a given situation. In any case, when the answer(s) was(were) not suitable to the circumstances, the SCANKEE system issues information about mistakes done by the person being trained. Somewhat more sophisticated possibilities are available for examining the skill, expertise and readiness of plant controllers to perform their duties. In this case a special mechanism, timing precisely all answers of the trained person, is set forward.

Conclusion & discussion

The developed system (SCANKEE) and the carefully tested method of its application in computer-assisted education of plant operators may be treated as a very general research and development tool to facilitate and speed up the educational process elsewhere. This finding was proven by a fact, that in any industrial organization operators must have their substitutes; moreover, industrial obobjects of large importance must be independent on typical probproblems issued by the staff itself, like opposition, incompetence, retirements, absence (justified, not justified), illnesses, etc.

References

Gutman, E., private information;
Hippe, Z.S., (1991). Artificial Intelligence in Chemistry: Structure Elucidation and Simulation of Organic Reactions, Elsevier (Amsterdam)/ PWN (Warsaw).

MULTI MEDIA COURSEWARE IN A LEARNING CONTEXT: PERSPECTIVES ON THE ROLES OF MEDIA AND HUMAN ACTORS

Robert Lewis, University of Lancaster, United Kingdom, & Neurope Lab, France; Jan Schoenmaker, ECC, Enschede, The Netherlands

Abstract
The role of resources to support learning are many and varied. This paper places human-human interaction at the heart of learning and discusses a variety of uses of information and communications technologies to support human learning. Traditional roles for the computer as surrogate teacher and as a mirror on learners' thinking are briefly reviewed. The new role for technology in bringing human experts into the workplace of professionals is described through recent and current projects concerned with learning and knowledge exchange. Media characteristics and current development models in educational technology are reviewed, as well as experiences with and perspectives on different types of applications.
Keywords: Computer-assisted learning; Open learning; Information and communications technologies; Vocational and professional development; Resource development; Multi media; Integrated Performance Support Systems

Roles for resources
The major preoccupation of a teacher is to make the best possible resources available through which students can be helped to learn. The list of possible resources is extensive and their selected use will be governed by the detailed goals of the learning process. Obvious items include laboratories with equipment, museums, art galleries, the natural environment, telescopes, texts, film, video, audio, computer technology, and so on. . . . . But if the goal is to help a novice to learn to ski, then a selection will be made from these but of prime importance will be skis and boots, and a slope with snow that is flat at the bottom, preferably prepared so that the task is made easier in the initial stages. Over and above all these resources however are human beings, either tutors or other learners. Whilst it is possible for a novice learner to find
out how to ski without human support, the task will be quite a lot more difficult and, moreover, it will not be half as much fun.
Many of the resources listed above can be used in different ways. For example, text in the shape of a book may aim to stimulate learners' understanding as they

500

read. On the other hand it may be a factual reference source, a set of challenging problems or instructions on how to operate a piece of equipment. The tutor will spend quite some time, not only selecting or creating the pieces of text for the learners but also thinking about the role that the text should have alongside other resources. As with text, so with all the other kinds of resource available to the tutor including human beings. The tutor's own role may be as a source of knowledge or it may be in support helping the learners get the most from other resources. Other learners may have a specific role in the collaborative solution to a set of problems. The successful tutor is the one who capitalises on the strengths of each resource.

The role of the computer

When the resources include educational technology, the role which the technology is expected to play is of prime importance. An analysis of some years ago illustrates (figure 1) how diverse this role may be, varying from surrogate tutor (on the left-hand branch) to interactive mirror on learners' ideas (towards the end of the right-hand branch).

"To avoid the serious confusion which frequently arises (in describing computer software use), it is important to state that in what follows the computer is nearly always being used to assist heuristic methods of education. Very little didactic instruction is involved and, specifically, the computer is not used to assess students' responses or performance"
(Lewis, 1980).

Once more the issue is of the appropriate role for the resource. This may be approached by asking a question about the strengths of alternative resources. Put another way, one may ask what is it that computers do well? and what it is that humans do well? The answer must include the fact that humans understand humans quite well and that computers are good at computation and searching. So we might conclude that it is inappropriate to give computers the role of assessing human understanding, but quite a good idea for them, for example, to perform complex calculations which are at the basis of a simulation. To this latter task could be added their ability to manipulate and display information in a range of alternative forms which are otherwise extremely time-consuming for humans.

In summary, learning goals may be met by:

- human actors like a teacher or a tutor and other students, in their own organizational and social context
- different media or resources like written materials, audio-video materials and software

In each learning situation a delicate balance has to be found in terms of the roles of these actors and resources. Special factors like teachers expertise and acceptance of resources will be influential.

Also , we must recognize (for the time being) instructional and learning theories offer only partial descriptions of the processes going on in a learning situation. This is due to the state of the art which is not advanced enough to describe the processes of instruction and learning at such a detailed level as is needed for the appropriate design of educational technology. materials.

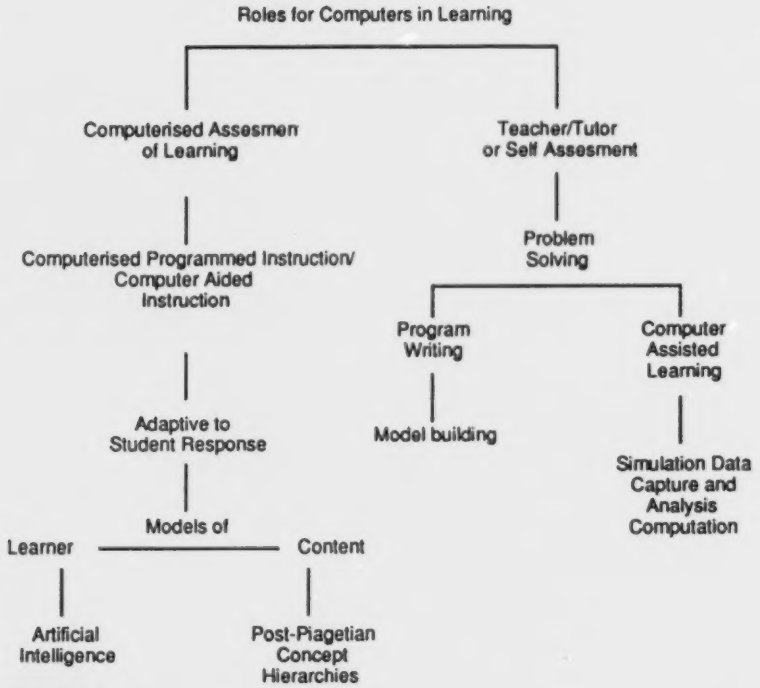


Figure 1
Roles for information technology in support of learning

Further, these instructional processes are subject to implicit and creative factors in the interaction between a learner and different human actors and are hard to model, certainly at a formal level.
The consequences of these statements are:

- the role of educational technology is context dependent, the human actors playing a decisive role
- designing educational technology -based products is not a straight forward matter based solely on sound theories, but is very much a matter of practical experience - based engineering as well as creativity and choice.

In order to further the analysis, an approach to teaching and learning by means of eductional technology will be presented from the viewpoints mentioned before, namely the human actors view and the resource based view.

The human actors view

Supporting human-human interaction

In recent years as communications technologies have developed, the role of technology can be even more diverse. A main goal in much current research is to explore how far communications technologies can be used to bring human beings together. In one type of application, technology now has the role of creating virtual classrooms in which tutors and learners may be located anywhere? This often depends on the use of direct broadcasting often using satellites to create synchronous (real-time) interaction. Asynchronous applications are being explored at Neurope Lab on the outskirts of Geneva. Before discussing these applications, it is necessary to examine the identified needs which the technology is attempting to address.

A few years ago a team at the University of Lancaster was asked by a government agency to investigate ways in which it might be possible:

"to improve the flexibility and responsiveness of vocational education and training provision in Great Britain through producing a plan for the creation of information technology based open learning provision alongside established practices in existing universities."

The resulting feasibility study produced a model of Information Technology-based Open Learning (ITOL). An outline may be found in the report by Hodgson, Lewis and McConnell (1989).

Before proceeding further, it is quite essential to remove any ambiguity from the term 'open-learning'. In this paper and in the models described, open-learning is quite distinct from distance education or distance learning in one specific and important way. In open-learning the curriculum is set by the learner. This is not the case in traditional distance learning courses where the learners follow a curriculum set by others.

The time, place and style of learning may be quite flexible in both learning scenarios. It is also useful to emphasise the distinction between the ITOL model and the various forms of campus-based uses of technology to support learning.

The outline of the ITOL model is shown in figure 2.

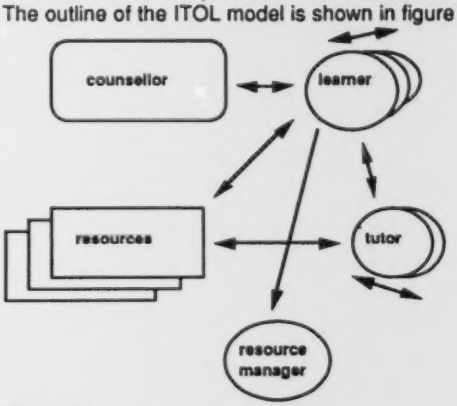


Figure 2
The ITOL Model

Fundamentally, it allows any individual learner to communicate with a tutor or tutors, or facilitator(s), with other learners and with a series of both university and non-university based resources. In addition, learners may have available to them a counsellor from outside the university system, most likely someone from inside their own organisation. Finally, there will need to be a resources manager with whom learners can also communicate.
Once more we should examine the role of the technology. In this case it is able to provide access to resources and support interactions of the kind impossible to do in other ways. One can communicate and get access to resources independent of time and place.
Asynchronous communications has enormous benefits: a major one is that it is non-invasive (how often does the phone ring at just the wrong moment?) and patient (how often do you phone someone and find that they are not there?). What is more, the respondent has time to think about the best reply to the question posed.

Professional development
A major project at Neurope Lab is one supported under DELTA in the CEC Third Framework Programme. The objective of the JITOL project is to develop and evaluate a learning environment for the professional development of individuals. The environment is based upon information-technology consisting of a telematic network linking and making accessible computer resources including data and knowledge bases, and using conferencing and
communications software to facilitate access to human tutors and co-learners. The project is outlined in a paper by Lewis, Goodyear and Boder (1992).
There are two aspects to the underlying rationale. Firstly, learning is basically a dynamic and continuous process.
Therefore, structured courses at fixed times do not correspond to the nature of learning needs. A continuously evolving, interactive open learning environment is more appropriate.
Secondly, today's state of knowledge is becoming increasingly complex, quantitatively difficult to comprehend and constantly changing. Therefore, it is more and more a challenge to select appropriate information for individuals to enhance their knowledge or to solve specific problems. In consequence, only knowledge bases which progressively integrate learners'
requirements stand a chance of meeting real needs. This is true both for learners at early stages of training and for the updating of highly-skilled professionals. This approach is termed "just-in-time knowledge transfer" to refer to the dynamic aspect of learning and to the appropriateness of knowledge provided to assist in solving actual problems. There are two specific foci within the JITOL environment:
the computer mediated human-human interaction process (HHIP):
In professional development contexts, the stimulus for a new learner action will often be a problem that arises in the process of the learner's work. A mentor or tutor and the learner will then engage in an iterative refinement (the human-human interaction process, HHIP) of their joint understanding of the learner's problem. Having established sufficient understanding of the learner's problem/question, the mentor may assemble a package of knowledge resources which the learner will study and then engage in further rounds of interaction with the mentor.

- the progressive construction of appropriate and evolving knowledge bases (EKB):
The idea behind the Evolving Knowledge Bases (EKB) is that representations of the interactions between participants in the system should be stored and thereby become additional knowledge resources for subsequent use. Since the interactions will mostly emerge from problems that arise in the course of learners' professional work, indexing of these knowledge bases can be problem-oriented. In this way, learners will benefit both from actual interactions on the network and from subsequent updating and customising of resources.
Experimentation with, and evaluation of the JITOL environment will be achieved through pilot investigations. After a nine month set-up period (which started in January 1992), there will be a first nine-month user trial where data will be collected in accordance with a project-wide evaluation, testing and validation methodology. Interactions between actors on the network will also be captured in order to enhance the knowledge bases for a further six-month user trial.

The resource or media based view

Introduction

In the resource based approach one focuses on the optimal usage of resources making use of selected media.
One can view a medium as an intermediary between an instructional and a learning agent in which two-way communication can be established.
Before the use of computers in education and training, one was talking about more or less passive media like written materials and audio- and video materials. These materials could support learning cycles in the sense of allocating a task, presenting information and presenting exercises and delivering answers. No one would say that these media are intelligent. With the introduction of computer-based learning adaptation of instructional strategies in terms of adaptive presentation and interaction became an issue in which intelligence seemed to be implemented. In this sense adaptation to a learners characteristics and foreknowledge is seen as a major aspect of 'intelligence'.
In practice not many 'intelligent' programs are available or are really being used. Their application in practice seems to be restricted, by the inadequacy of current learner models and the difficulties one has with formalization and generalization of instructional strategies. Also, the difficulty (perhaps impossibility) one has in modelling instructional strategies independent of the subject matter domain and learning context.
In discussions of different media one generally focuses on supporting cognitive processes while affective aspects of the medium are ignored. However, the combination of interaction and audio-visual support increases motivation strongly compared to other - less rich - media. Further, group work with computers has been shown to stimulate motivation. This is an important contextual factor which accepts the importance of the role of (other) human factors and also stresses the need to adapt the specific role of human actors and resources to a particular instructional situation.
In order to build up a model of the role of media we introduce the following characteristics of media (figure 3):
- the degree of interactivity with the learner (including adaptation to a learners profile)

- the degree of multi-sensory stimulus(not saying that more stimuli necessarily lead to better learning, but in terms of the bandwidth available)

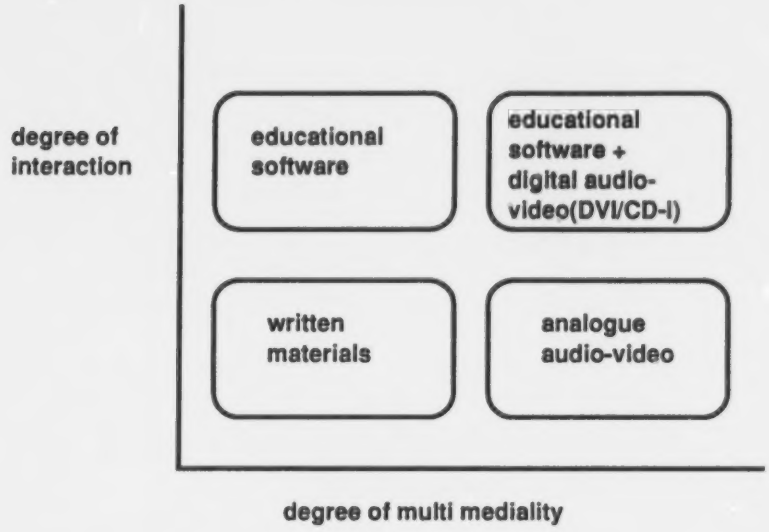


Figure 3
Media Characteristics

Media have to be viewed primarily from the functional point: what is the scope of the degree of multi-sensory stimuli they support and the degree of interactivity? The optical media move towards supporting a broad band high quality sound, stills and running video. This holds for video disc, CD-I and DVI. The degree of interaction possibilities however is different: videodisc only allows overlays, while the digital storage (if on a hard disk) of DVI allows run time manipulation of audio and video objects.
Storage capacity, performance and quality of sound and video are more technical features. In general they improve rapidly with new releases, without changing the basic media characteristics.
Other points in evaluating a technology are the distribution and updating aspects. These are major cost factors.
In introducing a new medium one should be aware that it might introduce a new and therefore yet unknown mode of interaction between user and application. This mode of expression has to follow its own evolutionary path. If one considers the evolution of the use of sound in films, one sees a gradual development in the meaning of sound. For instance a scene of suspense is supported by a certain type of music. With computers also the degree of interaction is a mode which has to find its evolutionary path: first time users - and still a number of developers I-

often expect a static presentation without interaction (viewing the medium as television) and a linear ordering of content.

Application types

One has seen an evolution of well known application types like drills, tutorials and simulations. In fact it would be better not to speak anymore in these terms because now applications are often mixtures of different types and have far more specific characteristics related, for instance, to the subject matter as in dictionaries supporting language programs.

Moreover, it is not longer the goal to put everything into the software part but only those features of the medium which are relevant to the educational or training need. An example is presenting an introduction in a topic by means of written materials, while developing skills could be supported by a simulation package, with assessment carried out by the teacher. So, each learning situation has to be defined and implemented taking account of contextual factors with the optimum use of the various media.

In business training a new situation has arisen which is influencing the whole training perspective and suggests a new vision. The experience in many training situations in business is that training which is separate from the working place may not be optimally effective. Although seen from a training perspective the use of computer-based training seems to be adequate, the broader perspective of performance (how employees performs their daily task) turns out to be quite different. If one takes 100% time allocated to acquiring necessary knowledge and acquiring skills in order to reach the necessary performance level, only 20% of the time (and money) is spent on regular training (Gery, 1991). The other 80% of time is spent in an ad-hoc way in the working place in looking back at the training materials, asking others for help, looking for manuals etc.

Therefore the need for integrating working and learning becomes an imperative. In the past the concept of embedded training seemed to be a solution. In the context of performance support training will be related to and integrated with other supporting activities like getting advice (from human actors or from expert system like advisors), on line help , and access to information resources. Under the name of Integrated Performance Support Systems (IPSS) new roles for human actors and resources produce a new perspective beyond training.

Also, the new technologies offer new opportunities. Hypermedia technology has been the basis of new applications in which user control is the key factor. Providing facilities for browsing through information sources with powerful support intervention has become a compelling instructional strategy.

As a generalization one could say that , in the past decade, applications have supported a wider spectrum of learner control and that the selection of resources by the learner has become more flexible resulting in better tuned learning environments. Applications offer options to select required modules, install all kind of supporting features like the degree of help and other variables such as the amount or repetition and review.

Development

The development of resources making use of media should be seen as an engineering activity.

This means :

- one makes engineering models on the basis of theories; primarily instructional and learning theories and software engineering principles (Schoenmaker, 1990)
- one should take into account implementation factors like the number of computers available, expertise and acceptance of teachers, organizational issues etc.
- one also takes into account the (im)possibilities of the medium itself. Here one should consider not only the hardware characteristics of the technology but also the system software characteristics, because they are decisive in terms of the functionality which can be implemented.

So engineering is an activity in which theoretical and practical aspects have to be combined and in which project management issues like time, budget and expertise available play a role. The high costs of educational software development certainly have impeded the acceptance of educational technology. Lack of formal and detailed methods, guidelines and experience and, of course, the relevance of aspects of choice and creativity which play a role have been important factors in this respect. One could say that educational software development is only slowly becoming an industrial activity.
The development process consists of analysis, design, implementation and maintenance. (plus 'use and guidance ' which has had to be omitted from this discussion). (A detailed analysis was undertaken during the DELTA Exploratory Action, - see Start -Up, 1991).What should be modelled is described by a product method, how it is to be modelled by a process method and the way in which it is controlled by a project management method.

Analysis and design

The design process is often not a straight forward matter, but very much a matter of making design decisions. In development one always has to find the optimum between freedom for the designers and standardization.
In educational software the major decisions concern instructional strategies, learner modelling and user interface design.
Different modelling approaches have been followed, for instance the instructional transaction approach. In this approach one distinguishes basic building blocks of educational meaningful interaction. A variety of types of interactions are possible between learner and computer, not only question and answer but asking for information by pointing to an object.
Different educational philosophies like behaviouristic, cognitive or humanistic approaches do not lead to simple models. Experience shows that these philosophies present a context which one can only model on an empirical basis. Experience also shows that one is dealing with an open system as new technologies lead to new instructional strategies. There is not one central theory supporting educational design. As such the design of educational technology - based materials with its demand for explicit and detailed models acts as a mirror for thinking about learning processes.
A conclusion might also be that design support should not be in the form of standards and fixed product models, but in terms of process support like reference models and cases as examples and guidelines. An example of an educational analysis and design support tool is COCOS (Rosendaal, 1992); this tool supports the modelling of objectives, tasks, activities etc., in fact of the characteristics of a

complete course- the curriculum analysis phase. The designer can ask for help in modelling objectives and strategies. However it remains up to the designer to decide about the way he wants to make use of this advise.

Implementation

A major discussion point in the past decades has been the adequacy of implementation tools. In general the debate was between using general programming languages versus authoring systems and authoring languages. Programming languages allowing maximum flexibility but less human efficiency in coding, while authoring environments as generators are more efficient but less flexible.

However the scope on development issues has to be broadened. Nowadays applications are so big that structured analysis and design (the first phases of development) and maintenance (the last phase) have become important. In general analysis and design take up 25-40% of total time and costs. Maintenance of programs can even exceed the initial development costs over a lifecycle of 3-5 years, although not many educational software programs have a long life cycle. So the search is for environments supporting not only implementation, but also analysis, design and maintenance. A demand arises then for making use of data dictionaries or repositories and automatic documentation. In general , CASE-tools (Computer Aided Software Engineering) and IPSE's (Integrated Project Support Environment) support this.

Nowadays general Userinterface Management Systems (i.e. Macintosh, MS-Windows, Presentation Manager, X-Windows) support on one hand functionality needed for educational software (most programs consist of more than 50% calls to these functions) and on the other hand support regular software engineering principles. Also portability of tools across these platforms is becoming more common. Additionaly, multi media data base tools are adequate for implementing and updating subject matter and user data. Existing authoring environments are rather primitive in the sense of structured storage and manipulation of data. Because of the increased size of applications this has become an issue.

For these reasons one might say that an authoring environment is not by definition more suitable than general tools.

In business the new perspective of Integrated Performance Support Systems brings opportunities to develop information systems and training and other functions with the same tools. By making use of general user-interface management systems and a multi media database management system with the opportunity to interface to a general programming language one create robust and flexible educational software. The advantage of making use of the same environment for developing information systems and for training and other functions, is that one can use the same resources during the development of them all..

Field testing

End users evaluation is not a popular matter in the world of educational software producers!. Although no one denies its importance, costs and especially the often long period involved in field testing are barriers. However, especially in educational software, testing is very important because:

- The end user is not directly involved in the analysis and design phase; instead a teacher or educational psychologist brings forward a model of the learners behaviour. This is different and more troublesome than in regular automation where one can analyse a real world situation which has to be automated. With educational software one often also introduces the computer as a new medium.
- As stated before it is not our viewpoint that current instructional and learning theories can deliver the models which prescribe in a complete and consistent way learning behaviour and the required instructional strategies.
- Contextual and local factors (amongst them social and organizational factors) play a major role.

One can also draw a conclusion with regard to research in the area of educational technology. Given these reasons for undertaking careful field testing one could also argue that in the case of educational research it might often be better to start with action research in practice in order to assess the variables involved. On the basis of such action research one could set up a fundamental research environment in such a way that one can study the wanted variables and eliminate others.

Concluding remarks

Whilst technology has had some success in supporting traditional learning through computer-based learning in schools and computer-based training in industry, the potential contribution of advanced communications has hardly been tapped.

Rather than follow the earlier pathways which some applications trod of trying to usurp the role of human mentors in the learning process, new approaches build upon human skills in personal communication. At the moment computer-mediated human-human communication channels are impoverished even though some exciting elements can be demonstrated. The challenge is to integrate the best of these into coherent convivial systems.

The open learning support environment described needs access to rich resources of many kinds as well as to humans Amongst these resources are those which support learning in new ways which are more flexible and capable of being tuned to specific user and organizational needs. Given the size and complexity of new educational software and the need to maintain and update such resources, the stress is more and more placed on structured analysis and design rather than on implementation. Advanced software engineering tools become indispensible in meeting present day demands and these are taking over from dedicated authoring environments. A merging of general and dedicated tools may bring the strengths of each together. The digitization of video and audio is making multi-media databases of resources completely manipulable by software.. However the semantic descriptions of the integration needs and processes are not yet well defined. New concepts such as Integrated Performance Support Systems create new perspectives, challenges and opportunities and may provide a new common

denominator for the integration of information systems applications and the corresponding training andsupport software.

References

Gery, G, (1991). Electronic Performance Support Systems. Weingarten Publications, Boston

Hodgson, V, Lewis, R., & McConnell, D. (1989). Information Technology-based Open Learning. Occasional Paper InTer/12/98, ESRC-InTER Programme, University of Lancaster

Lewis, R., Goodyear, P., & Boder, A. (1992). Just-in-time Open Learning - a project outline. Occasional Paper NL/1/92, Neurope Lab, Archamps.

Lewis, R. , & Want, D. (1980). Educational Computing at Chelsea (1969-79). In: ,Computer Assisted Learning, eds. R. Lewis and E.D. Tagg). p. 163-175. Heinemann Educational Books, London.

Rosendaal, B. (1992), COCOS, A Workbench for trainers in organisations. ECC-report, ECC, Enschede

Schoenmaker, J, Nienhuis,E, Scholten, J, & Titulaer, J, (1990). A methodology for educational software engineering. Computers in Education, p.189-194, Elsevier Science Publishers B.V., Amsterdam

START-UP (1991). Production of Multimedia Materials for Education and Training in Europe: Start-up Report. Available from OTE,133 rue de Sevres, F-75006 Paris

CURRENT TRENDS IN EDUCATIONAL TECHNOLOGY: A METHODOLOGY AND FINDINGS

D. P. Ely, Development and Evaluation and ERIC Clearinghouse on Information Resources, Syracuse University, USA

There are many ways in which trends could be identified: expert opinion, panels of specialists or informed observation. This study chose content analysis as the primary vehicle for determining trends based on earlier works of Naisbitt (1982) and his model (Janowitz, 1976). The basic premise of these works is that current trends can best be determined by what people are saying publicly, through newspapers and magazines. This study, and those that preceded it, used the same basic procedure: identification of emerging topics in key publications over a period of one year. The rationale seems sound. It is possible to determine trends by reviewing what people are saying publicly about matters within the field. The literature of the field was selected as the best comprehensive coverage of current thinking and events in the field. A selected body of literature was carefully reviewed using a team of educational technology specialists to help determine the status of the field today.

Literature Sources

To maintain consistency from year to year, the same sources of information were used as in the 1988 and 1989 trend studies with a few exceptions. To aid in the selection of sources, the Moore and Braden (1988) report on "Prestige and Influence in the Field of Educational Technology" was used. The people, publications, and institutions of "high prestige" were identified by a survey of personnel in the field. This trends study used the highest ranking journals and the dissertations produced by the top universities as major sources of literature. Additional sources of data included the papers given at major national and international conferences and input to the ERIC database in the field of educational technology. Conferences are one of the most visible ways of presenting new ideas and findings to colleagues and therefore contribute to the trends in the field. The ERIC system solicits unpublished materials such as reports, evaluations, studies and papers to review and, following evaluative criteria, to select for inclusion in the database. The journal sources were: the British Journal of Educational Technology, Education and Training Technology International, Educational Technology, Educational Technology Research and Development, and TechTrends. Dissertation sources were: Arizona State University, Florida State University, Indiana University, Syracuse University and the University of Southern

California. Conferences reviewed were: Association for Educational Communications and Technology, Educational Technology International Conference and the National Society for Performance and Instruction. The dates of journals, dissertations and ERIC input were 1 October 1990 - 30 September 1991 and the conferences were all held in 1991.

Content Analysis

Some concern about the use of content analysis centred on three factors:

- analysis of large bodies of text;
- the validity and reliability of coding;
- the translation of quantitative data to descriptive trends.

These concerns are addressed in the larger paper.

Trends

Content analysis of the British and North American literature reveals the following trends:

- The creation of technology-based teaching/learning products is based largely upon instructional design and development principles. • Evaluation has taken on greater importance as the concept of performance technology has been further developed.
- The number of educational technology case studies is growing and provides general guidance for potential users.
- Distance education is evident at almost every educational level in almost every sector.
- The field of educational technology has more and better information about itself than every before.
- Computers are pervasive in the schools. Virtually every school in the United States has computers.
- Telecommunications is the link that is connecting education to the world.
- The teacher's role in the teaching and learning process is changing as new technologies are introduced into the classroom.
- There is increasing pressure for the schools to consider the adoption of technology while, at the same time, concern is expressed for the impact of technology on children in the society at large.
- Professional education of educational technologists has stabilized in size and scope.

Trends 1988-1991-- A Longer View

Using studies conducted in 1988 and 1989 as baselines, the following longer range trends have emerged:

- Educational technology is being shaped more by external forces than by the internal influence of its own professionals.
- The use of traditional media resources has become routine in most elementary and secondary classrooms.
- There is little evidence to show that the computer has made major contributions to learning in the classroom other than to help learners know how to use it.
- The self-contained classroom is the greatest single barrier to use of educational technology principles and practices.

- The field is shifting from the use of media and technology for enrichment to technology for replacement.
- Instructional development is being practiced more in non-school settings than in schools.
- Distance education has become an operational analog of educational technology.
- Cognitive science provides the best source of theoretical principles that underlie instructional design.
- Evaluation is valued but infrequently used.
- Educational technology continues to be perceived as a field concerned more with hardware and software than with its applications for teaching and learning.

The trends stated here are current and are not necessarily predictive. They provide indicators of developments within the field but do not determine the future. They come primarily from the literature of Canada, the United Kingdom and the United States and should be considered in those settings.

References

Ely, D. P. (1989). Trends and issues in educational technology 1988. Syracuse, NY: ERIC Clearinghouse on Information Resources.

Ely, D. P. (1990). Trends and issues in educational technology 1989. Syracuse, NY: ERIC Clearinghouse on Information Resources.

Janowitz, M. (1976). Content analysis and the study of sociopolitical change. Journal of Communication, 26(4), 20-21.

Naisbitt, J. (1982). Megatrends. New York: Warner Books.

ON THE INTRODUCTION OF COMPUTERS INTO THE TURKISH EDUCATIONAL SYSTEM

Yasar Ersoy, Middle East Technical University, Ankara, Turkey

Abstract

The general question addressed in this study was to describe the introduction of computers into the Turkish educational system (TES) and to explain the main problems and current issues related to the implementation of computers in classrooms and the school teacher education very briefly. This short study indicates that there are plenty of issues and constraints in the classroom practice with computers such as the need for the use of generic packages and the continuous training of school teachers.

Introduction

New information technology (NIT), in particular computers, is a powerful tool for the learner and teacher/instructor since it creates various possibilities for organizing the communication between man and machines, and enriches the teaching/learning process to a great extent. However, capabilities for computerization differ greatly between societies in the developed and developing countries. In this paper, we will attempt to overview the present situation of introducing micro-computers into the Turkish educational system (TES) and describe the main issues related to the Computer Based Education (CBE). Computers have been introduced to the TES for at least two reasons. Firstly because NIT surrounds us, i.e. the workplace uses it in many areas, and the school must reflect this. Secondly, because NIT adds new qualitative and quantitative possibilities to education and thus enables it to meet the demands for qualification. After some experiences in different schools it is realized that there is an urgent need for the use of generic packages, development of adequate software and teaching/learning materials and the continuing education and support of school teachers. Moreover, it seems that the school managers' view, the teaching habits of teachers and their attitudes toward computers and CBE may be a considerable barrier to set up new teaching/ learning environment and improve the quality and level of general education in developing countries.

Computer-related activities in school

The Turkish Government has put some effort into the introduction of computers into the TES since 1985 but on a small scale as a learning tool. The ME is responsible

for the purchasing hardware, the organization of training courses for teachers, the evaluation, validation and distribution of educational software (SW) to the schools.

Preliminary and Pre-Pilot Study Periods (1984-87/89)
During the preliminary period (1984-87) and pre-pilot study period (1987-89), there was a growing awareness a lot of projects were going on over the world. Some pressure from parents and private computer firms for the introduction of computers into schools were also observed.

Preliminary Period (1984-87)
During this period, there were various unspecified and unconnected explorations with computers and learning in Turkey. There was a general motivation that computers would be a part of future, and that Turkish children should be prepared for this new environment. Various public and private schools (primary, secondary and vocational schools) purchased and accumulated different types of incompatible home computers for playing games and a few microcomputers. This lead to the situation in 1987 of approximately 3000 computers in 400 schools in the country. The training programme for teachers was designed by the computer firms and differed widely in duration and content. Teachers were being familiarised with hardware and sometimes with programming languages, in particular BASIC.

Pre-pilot Study Period/ Initiation Phase (1987-89)
During this period the one of the Minister of State initiated the introduction of computers into schools and the ME was not directly involved in the activities at this point. Then the ME was directly involved in the use of the NIT in education and brought some experts in a working group of the General Assembly of National Education, [1]. A special committee of experts was also formed in the State Planning Office to prepare a special report on the information technology and education for the development of the country, [2].

Experimental Phases I and II (1989-90/91)
Experimental Phase I (1989-90)
This was an initial and starting activity for the first experimental phase of the announced big project. This experience can also be considered as the pre-pilot project phase since various national and international private computer firms, companies or agents were invited to install a computer laboratory of 20 computers in a group of selected schools in Ankara, Istanbul and Izmir. The school subjects chosen to teach and develop software were usually mathematical sciences and a few technical subjects. By June 1989 the ME appointed an evaluation committee to review the computer laboratories set up by the private firms and the associated software, and organized a conference in June 1990 to get the views of specialists and the response of general managers and the executors of the Project, [3].

Experimental Phase II (1990-91)
The main goal was faster the installation of computers in schools and to gain some experience in different areas of study of the introduction of computers in the TES as soon as possible. Based upon the policy of regional equity, it was decided that each of 73 provinces would have at least two computer laboratories during the

second phase of the experimental period. Many different school sectors of the TES and various school subjects were considered to fulfil the requests and achieve the goal partly. During this period 6500 computers were purchased for about 300 schools, 5000 hours of educational SW for approxi-mately 100 courses were developed and 2400 teachers were trained, [3].

General assessment and critics of the project

What is the place of CBE in the TES? The question is a simple one to pose, but it calls for an immense task of enquiry to answer. This is because the world "place" invokes questions of educational policy, availability of experienced teachers, availability of resources, issues of content, issues of method, implications of the need and etc. All these issues had been discussed in several national committees during the pre-pilot period and the experimental phases, and approved by the ME and the Government.

Almost all SW developed in the experimental phases of the Project do not yet sufficiently exploit the capacity of the computer to enhance teaching and learning. The following are some characteristic features of the developed SW.

- *Tutorial Mode of CBE:* Almost all of the software available are in the tutorial mode of CBE and do not generally motivate the student to learn by himself.
- *Emphasis on Traditional Approaches:* No clear presentation of the process of problem-solving in the developed SW exists.
- *Great Influence of Textbooks:* Another characteristic feature of the situation is the dominance of the published textbooks on the educational SW.

The Inservice Department of the ME has given some projects to several universities relating to in-service training of teachers. In the experimental phases and before, almost all universities generally interpreted the requests in the projects by supplying courses on how to use a computer from a technical perspective with some additional time spend on BASIC programming.

Main problems and issues

- *Ignorance in Scientific Education:* Some senior officers, school teachers and most people ignore the scientific education, pedagogical problems and issues in schools. They should be aware of the developments and new trends.
- *Modification and Integration of the Curricula:* The curricula as whole needs to be revised, modified and integrated to reflect the changing needs of the information society. Therefore the general aims and goals of compulsory education and of others should be overviewed.
- *Reorganization of Teaching/Learning Environment:* Most of the courses at schools in Turkey are far from the ideal learning/teaching environment and efficient infrastructure. Almost all courses at present time are lecture-based far from motivating the student etc.
- *Changes in Teaching/Learning Process:* The teaching process in the school is still group- oriented because of the very large number of students in the class. There is no attempt or intention to change with the student-centred and individualized process.
- *School Managers' and Teachers' Attitude:* As emphasized before, one of the main factors and/or criteria for the post experimental phase is teachers' attitude toward C/CBI and their qualification in the educational practice. It seems that almost all

teachers in Turkey accept the new technology, in particular computers, in the educational environment if it is a real support to them in their classroom tasks, adequate training is provided and good material is supplied.

- *Expertise and Human Resources:* Another current problem is the lack of widespread knowledge, experience and skills in developing SW and training teachers. For example, few universities are in a position to provide even one course of programming and/or CBE for prospective teachers and practising teachers.
- *Need for Free Tool-Box:* Teachers develop their own programmes. Financial support may be provided to some pre- evaluated projects on SW development.

Concluding remarks

Besides reorganization and fundamental changes in the TES, there is an urgent need for the use of generic packages, development of adequate software and teaching/learning materials and the continuing education and support of school teachers. Moreover, it seems that the school managers' view, the teaching habits of teachers and their attitudes toward computers and CBE may be a considerable barrier to set up new teaching/learning environment and improve the quality and level of general education in developing countries. Therefore a special attention should be paid to the pre-service education and continuing training of teachers in such countries before introducing computers into primary and secondary public schools.

References

12. Milli Egitim Surasi Raporu; MEB Yay., Ankara, 1989.
VI. Bes Yillik Kalkinma Plani, Bilgi Teknolojisi ve Egitimi Ihtisas Komisyonu Raporu; TC Basbakanlik-DPT Yay., Ankara, 1990.
Türkiyede Bilgisayar Destekli Egitim; MEB-METARGEM Yay., Ankara, 1991.

WHY TEACHERS TEACH CONTENT AND CONTEXT AREA TOPICS OF COMPUTING

Klaus-Henning Hansen, IPN an der Universität Kiel, Germany

Computer education courses are well established in most German schools. Students are expected to learn principles of hard- and software structures, how to use data bases and text processors, and social aspects of this technology. These goals are reflected by curricula containing content, application, and context area topics. However, teachers who most frequently use computers in their lessons are rather seldom prepared for teaching social, historical, or ethical issues of computing. In this paper I will ask for the teachers' preferences to teach different types of topics and for causal relationships between teaching practices, teachers' characteristics, and the computer education curriculum.

Innovation theory offers models explaining why teachers change traditional teaching practices (Fullan, 1982, p. 107-129). However, little is known about individual dispositions that make teachers overcome the problems of a new and many- layered subject. Without substantial theory it seems useful to inspect empirical data guided by some general hypotheses. The data at hand came from a random sample of 1121 lower and upper secondary education teachers selected from 999 German schools. They were collected in 1989 as a part of the German COMPED-study conducted by the International Association for the Evaluation of Educational Achievement (IEA) (Pelgrum & Plomp, 1991) together with the Institut für die Pädagogik der Naturwissenschaften (IPN) (Hansen & Lang, 1992).

Path analysis based on linear regression was carried out on a subsample of 485 male and 56 female computer using teachers. Teaching preferences were measured by 26 items grouped under the headings "Computers and society", "Applications", "Problem solving and programming", and "Principles of hard- and software structures" asking for the topics teachers taught in their classes. The same items were rated in terms of the topics they learned during teacher education. Other variables selected for the analysis are the teachers' motivation to learn more about computers, a self rating of their computer science knowledge (including abilities necessary to work with computers), gender, and school level. Ordinal measurement of most variables and a reduced sample did not allow to test the structure of the path model *and* the measurement of the items by programs like LISREL (Jöreskog & Sörbom 1988). Hence the quality of the scale items was controlled separately for each scale by calculating alpha reliability coefficients. The construct validity of the items was tested by selecting compact and discrete clusters from multidimensional scaling solutions. This method was producing inde-

pendent subscales for computer science topics and context/application area topics, either taught in classroom or learned during teacher education.
Though no explicit path model was specified in advance, the analysis was guided by some general hypotheses:

- Female teachers more frequently than male teachers look beyond programming or hardware in their computer education courses;
- Both teachers' characteristics (attitudes, interests, prior knowledge) and the curricula determine which kinds of topics teachers teach in classroom;
- Teachers usually do not teach about computers until they feel qualified in a field where many students have acquired a high degree of specialized knowledge;
- Their self concept can be reinforced by teacher training yet gender differences or differences based on individual experience or motivation also have an effect;
- Context area topics are learned or taught if computer science topics are learned or taught but not vice versa.

Path analysis allows to distinguish direct and indirect effects of the variables involved. It helps to decide, for instance, whether the teachers' motivation to learn about computer related topics directly encourages new teaching practices or whether it makes in service courses attractive in the first place. Similarly we may ask whether teacher education directly changes teaching behavior or enhances the self concept of teachers concerning computer science and/or context area topics. While a measure of self ratings for computer science knowledge/ability was available, no such items where at hand for knowledge about context area topics.
The assumptions listed above were used to develop a first path model based on 10 variables. The model was tested by linear regression and a simplified version of LISREL (Jöreskog & Sörbom 1987). Standardized residuals and test statistics were used to modify prior structural assumptions. The final model uses only 8 variables. It still has two path coefficients with a t -value below 1.96, however the overall fit is sufficient.

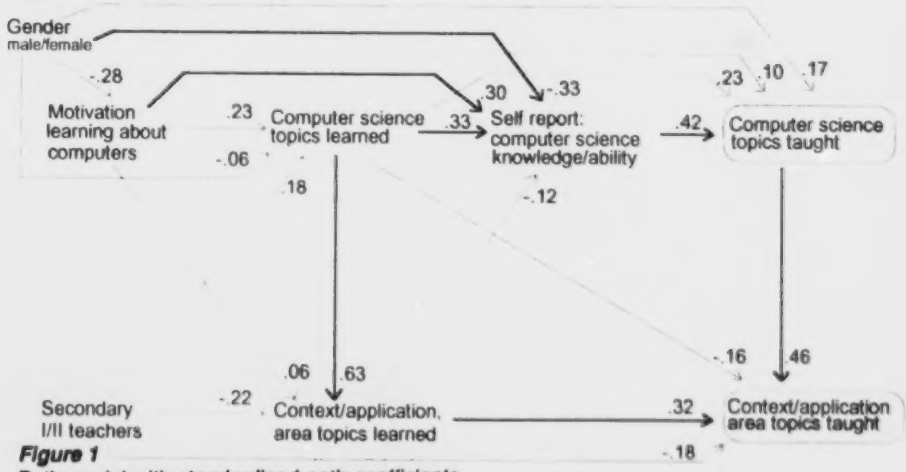


Figure 1
Path model with standardized path coefficients

Altogether figure 1 does not confirm the assumption that gender directly affects teaching preferences. Instead, a negative effect of gender is directed to motivation and self reported computer science knowledge. When interpreting these coefficients, we have to take into account the skewed gender distribution. Figure 1 also indicates that secondary II teachers less frequently teach context area topics than secondary I level teachers. This can be explained by an integration of context area topics into secondary I level curricula in the Federal Republic of Germany (Hansen, 1991).

Not less remarkable is the effect of teacher education on teaching practices compared with other variables. The impact of the curriculum was not measured directly but represented by the school level (secondary I/II teachers) and by hypothetical arrows from computer science topics to context area topics. Both have rather high path coefficients of .63 and .46 respectively. Though path analysis cannot strictly proof causality, the model does not reject the hypothesis that context area topics are taught if computer science topics are taught but not vice versa. Moreover, significant path coefficients between school level, topics learned during teacher training, and computer science topics taught in classroom support the assumption that the curriculum determines teaching practices. Both types of topics are more frequently taught if teachers have learned about them during teacher training. However, a positive self concept of their computer knowledge seems to come between teacher education and teaching practices.

A discussion of the results has to refer to the limits of the sample and to measurement problems. The exclusion of teachers not using computers reduced the sample to 541. The reliability of the motivation scale is below .8 and the construct validity of the all the scales was checked only by selecting clusters within a selection of separate scale items. Another important problem is posed by the specification of the path model. The path from computer science topics to context area topics is based on weak measurement and the model does not allow to test a feedback relation. However, the model clarifies the effect of in service training on teaching practices and it can be used to test data from other nations taking part in the IEA COMPED study.

References

Fullan, M (1982). The Meaning of Educational Change. Toronto: OISE Press.

Hansen, K.-H. (1991). Integration gesellschaftlicher und geschichtlicher Lernziele in das Informatik-Curriculum. In: G. Cyranek , H.J. Forneck, H. Goorhuis (eds.). Informatik-Curricula und Lehrerbildung. Frankfurt: Moritz Diesterweg (pp. 75-92).

Hansen, K.-H., & Lang, M. (1992). Computer im Unterricht. Ergebnisse der deutschen IEA-Studie. IPN Kiel, in press.

Jöreskog, K.G., & Sörbom, D. (1987). SIMPLIS. Mooresville. Scientific Software Inc.

Jöreskog, K.G., & Sörbom, D. (1988). LISREL VII: A Guide to the Program and Applications. Chicago, SPSS Inc.

Pelgrum, W.J., & Plomp, Tj. (1991). The Use of Computers in Education Worldwide. Results from the IEA 'Computers in Education' Survey. Oxford: Pergamon Press.

COMPUTER SUPPORT OF OPERATOR TRAINING BASED ON AN INSTRUCTION THEORY ABOUT PARALLELISM

F.B.M. Min, P.G. van Schaick Zillesen, & M.R. Gmelich Meijling,
Department of Education, University of Twente, Enschede,
The Netherlands

Summary
At the department of Education at the University of Twente, Enschede, The Netherlands some universal design systems have been developed in the period 1984-1989, with which a trained courseware designer can make computer simulation programs in the fields of physics, biology, medical science, and economics. Technical installations, processes or plants are also excellent subjects for this type of simulation. The educational computer simulation programs created with these systems, are characterised by a highly graphic interface. One of the design systems which was used was MacTHESIS and the design method which was used to develop computer simulation programs with this system was the so called 'MacTHESIS philosophy'. For the chemical company AKZO Hengelo, we developed a computer simulation program named BRINE PURIFICATION, based on this philosophy. These investigations led to a theory about parallelism.

Introduction
The simulation programs made with the MacTHESIS system link up completely with the Apple desktop publishing philosophy. The simulation programs themselves look like loose sheets of paper, which can be moved by means of the mouse. On the screen certain interesting visualized 'dynamic' processes take place. In this case a program is made about a complete chemical plant of the brine purification of the local mining industry. In the dynamic processes on the screen can the trainee can intervene by clicking the mouse in special, so-called 'in-click regions' or else just with the help of pull-down menus. The computer simulation program BRINE PURIFICATION is tested on the plant. A stand-alone version, is being tested only with coached paper materials, and also a version with an integrated tutorial instruction shell built in HyperCard is tested. This type of learning environment, in which CBT-materials are used besides an educational computer simulation program, is developed at the University of Twente and is called: an 'integrated computer simulation support learning environment', by Van Schaick Zillesen.

The MacTHESIS philosophy and some instruction methods

The MacTHESIS philosophy behind the computer simulation programs of Min et al. is described and showed in products on many conferences. The design environment, the teacher environment, and the learning environment, each have their own characteristics design parameters and a simple human-computer interaction.

From this point of view we developed and researched various instruction methods for use parallel to our simulation programs. Even the vizualized conceptual mathematical model plays an important, special role in the instruction in the learning environment. We mark 6 important instruction methods for parallel use:

- a student manual (beside the program);
- loose papers and worksheets (beside the program);
- a (interactive) highly visualized conceptual scheme of the model (in a parallel window);
- a help system (in a parallel window);
- 'video messages' as intelligent feedback ('desktop video'; in a parallel window);
- an instruction program ('CBT materials'; as a parallel proces).

In a learning environment for computer simulation it is essential that the learner or trainee can see both the simulated world and the instruction materials. Because of this requirement it is necessary to furnish the computer simulation program in such a way that important information remains in view.

The earliest simulations had no more than a graphic output on a screen, the instruction in relation to the computer simulation program was given in a separate textbook with or without loose worksheets.

MacTHESIS, a computer simulation system developed at the University of Twente, partly solved this problem of paper materials by creating an extra window, parallel with the program on which a survey appeared of the visualized, underlying mathematical model, allowing for intervention at the same time. This rather simplified form of instruction required other paper materials as well.

Research

We have tested a lot of hypotheses. One of the most important hypotheses being tested was: if parallel represented processes in windows are more effective than sequentially represented processes. The program BRINE PURIFICATION / PEKEL was used in that research. For the evaluation of this program a group of 9 employees from AKZO-Zout Chemie Hengelo were available. All subjects possessed advance knowledge about the structure of the simulated system (the brine purification installation) and about the relations within this system. One subject (process operator) was experienced in controlling the installation in practice, the other subjects had no practical experience in this respect. None of the subjects had ever worked with a program made by means of HyperCard and MacTHESIS software before. During the computer sessions the subjects were observed. Furthermore their interactions with the program (the simulation and the tutorial) were logged. We have proved that MacTHESIS is a good system for designing computer simulation program for education and training and in this case for a computer simulation program in color about brine purification. We developed a prototype in which information was transferred from the simulation program to the stack, enabling instructional feedback after the termination of the execution of

the simulation program, and vice versa, enabling the case and screen layout of the simulation program to be set from the HyperCard stack. We found that prototypes for CBT-materials can be developed in an extremely fast and highly interactive way using HyperCard. The philosophy behind the products made by this design method and there products were clear and useful in training situations. The general impressions of the subjects to the prototype were qualified as 'clear', 'instructive', 'pleasant' and 'interesting'. Furthermore according to he testees, the prototype can be used as a stand-alone learning tool and the guidance during the computer session is adequate. The subjects agree on the point that it was pleasant not to have separate paper materials. However, the statement that paper materials are more practical has not been rejected

The Parallel Instruction theory

It turned out that for a large number of pupils and trainees everything in a learning environment should be within reach and crystal clear. This was already well known for simulations, but technically hard to realize by courseware writers or most of the educational software houses. With tutorial CBT materials it is known that certain pupils are annoyed by the disappearance of the subject matter which has been read. Technically speaking calling back text which has been read is for almost all tutorials still problematic or not common sence.
The problem with separate paper materials is that students and trainees think they can do without. Even teachers tend to think so. As a result beautiful computer simulations remain unused at schools, in spite of the perfect design, styling and user-interface. Our earlier learning environments are characterized by modern input and output techniques and a wide range of different kinds of visualizations, from abstract to concrete. In spite of all that it turned out that the instruction method was the decisive factor. MultiMedia and HyperText techniques of in particular Macintosh computers, and the MacTHESIS system and philosophy, made it possible to prove our theories.

Conclusions

The investigations with the MacTHESIS philosophy and our designing method has resulted in a new MacTHESIS system, version 5.0 with desk-top video output and in an theory about instruction, the PI theory. The design method based on this instruction theory for computersimulation we have called the 'Parallel Instructional Theory for Simulation' design method, the PITS design method. Students of the University of Twente can follow three courses in which they learn the 'PI theory', the 'MacTHESIS philosophy' and the 'PITS design method'. More than 30 experimental programs are made until 1992. So the basis of our ideas how to built new simulation environments, based on our methods, techniques and this instruction theory about papallelism, is large.
The design and some learning environments will be published on CD-ROM on the end of 1992.

COMPUTER SUPPORT OF PROCESS ENGINEERING EDUCATION

Pieter G. van Schaick Zillesen, Marcel H. Zwietering, & Klaas van 't Riet, Department of Food Science, Wageningen Agricultural University, The Netherlands

This paper focuses on the instrumentation and implementation of computer aided instruction for process engineering education. The results of field–tests of two *simple–structured, computer–controlled, textbased tutorials* (balance and calculate) and two *graphics–based, student–controlled simulations* (tempe and yogurt) are reported. It is concluded that tutorials can be used to support the teaching of basic process engineering skills such as the skill to calculate mass and energy balances and the skill to estimate the order of magnitude of variables. Simulations can be used to support the teaching of higher order process engineering skills such as the scientific approach to problem solving, system analyses and predictive modelling.

Introduction

In recent years educational software has been introduced in the process engineering curriculum of many universities to serve purposes such as:

- In the early phases of process engineering education students have to learn basic skills (e.g., calculating energy and mass balances and estimating the order of magnitude of variables), which have to be applied in the later components of the curriculum. Time needed to complete the later parts may increase dramatically, if the students do not properly master these skills. Time consuming training sessions are usually implemented in the curriculum to enable the students to master these skills. During these sessions the professor presents exercises, checks the solutions of the students and gives instructional feedback (e.g., hints or corrections). Computer programs may be powerful tools for optimizing these training sessions, since most of the didactic activities mentioned above can easily be computer programmed.
- Traditionally, practical work is considered a very important component of the process engineering curriculum. Until recently practical work has been the best method to reach many important didactic goals such as the enhancement the students insight and the training of the scientific approach to problem solving. The achievement of these goals requires personal experience obtained by means of active experimentation. This active experimentation involves several phases. The main phases are: the systematic analyses of a problem, the design of hypotheses, the design of experiments for testing these hypotheses and the interpretation of the

results of these experiments. Furthermore, practical work can be used to confront the students with the abstract concepts introduced during lectures in a less abstract way. It also may be used to improve the students' appreciation of process engineering. However, often these goals also can be achieved by means of sessions with educational computer simulations. The decision to implement a computer simulation instead of (or besides) practical work depends on many factors such as: costs, safety, prerequisite skills and ethic factors.

- The use of software by professional process engineers is rapidly increasing. Apart from standard software (e.g., wordprocessors and spread sheets) process engineers use special purpose software for designing, predicting, optimizing and managing processes. Educational computer programs can be used for teaching the students to use this special purpose software.

Until now, the quality of most of the computer software used for teaching process engineering has been poor. This poor quality may be caused by a lack of knowledge about the instrumentation and implementation of this type of software. We designed, constructed, implemented and field–tested several computer programs to decrease this lack of knowledge. In the first phase of our research we field-tested two educational simulations (tempe and yogurt) and two tutorials (balance and calculate). The aim of the field–test was to evaluate the following questions:

- Are students able to learn basic process engineering skills by means of tutorials?
- Does the time needed to instruct the students decrease by the replacement of traditional training sessions by computer sessions with tutorials such as balance and calculate?
- Which type of training session do the students prefer: a conventional session or a computer session?
- Which type of session does the professor prefer: a conventional session or a computer session?
- Are students able to learn modelling skills by means of simulations?
- Should simulations such as tempe and yogurt be applied in a purely open context or should the students be guided (e.g., by means of instructions provided on paper)?
- Do students like to use simulations besides real experiments?
- Does the professor like to apply simulations besides real experiments?

Method

Students

We tested the prototypes during a university course (fermentation of food). The course made part of the Food Science and Engineering curriculum. Students attend this course during the third year of their regular program (Food Science and Engineering). The software was field-tested for the first time in 1991 (21 participants). The field-test was repeated during the next course in 1992 (25 participants). During the second course improved versions of the prototypes were tested.

Balance

The tutorial balance teaches and trains the procedure to calculate elemental balances of fermentation processes. Furthermore, balance trains the skill to

estimate the order of magnitude of variables. Balance consists of 27 exercises. At the start of each exercise the computer presents information to the student. In all exercises the student is requested to calculate the value of a variable. If the student's answer is not correct the computer presents the procedure to calculate the variable and requests the student to use this procedure to calculate the variable. If the student's answer is still not correct, the correct answer is presented to him. In 15 exercises the students are requested to estimate the value of the variable before calculating it. Based on the student's estimate a line of feedback is given.

Calculate
The tutorial calculate teaches and trains the procedure to calculate the balances for heat, biomass and oxygen of a solid state fermentation process. The structure of the student–computer interaction in calculate is the same as that described for balance. In an earlier phase of their studies the students studied the fermentation of tempe from blocks of soya beans by means of practical work. Calculate uses the same fermentation process as a case for making the calculations.

Tempe
Tempe simulates profiles of temperature, oxygen and biomass in a block of tempe. Tempe and calculate both are based on the same model equations. The goal of tempe is to teach the students the skill to predict the outcome of a process by means of computer simulation. Students can analyze the dynamic aspects of the simulation by changing the conditions under which the simulated process operates and observing the results. Furthermore, tempe can be used to teach the skill to make sensitivity analyses of a mathematical model by systematically changing the values of its parameters. The student completely controls the student-computer interaction. During the first field–test tempe was implemented in an open context. During the second field–test, the students used the program in combination with suggestions for simulated experiments, provided on paper.

Yogurt
Yogurt simulates a yogurt fermentation process. In yogurt two micro–organisms grow together in symbiosis. The interactions between the micro–organisms, the substrate and the product are complex. Yogurt and tempe share many characteristics; both simulations are used for teaching predictive modelling and both simulations are completely student–controlled. The students used yogurt in combination with paper materials with information about the mathematical model on which yogurt is based. Furthermore, the paper materials were used for suggesting simulation experiments.

Procedure
Four computers were available. On each computer one of the programs was implemented. During the course all students worked in pairs. All students worked with all of the programs (but in different order). The interactions of the students with the computer programs were logged automatically. The students were allowed to work with each program for maximal four hours. At the end of each computer session the students filled in an evaluation form.

Results

Most students were able to control the programs without any assistance. In case of problems the professor demonstrated the main control functions of the program. No problems referring to program control were observed after these demonstrations. The students reported on the evaluation forms that all of the programs were easy to control and pleasant to work with. Furthermore, the students reported that the goals of the programs were clear.

The students seemed to take great interest in the information presented by the computer programs. They reported on the evaluation forms that the information was clear and that there was enough information. During the computer sessions they frequently discussed the observed information with their partners. The students did most exercises without any assistance. Analyses of the logfiles showed that the questions in the tutorials were not too difficult to the students, since only a few wrong responses were recorded. Analyses of the logfiles of tempe showed that most students had been experimenting in a very systematic way.

The professor helped the students, when they asked for it. However, help was not often required. The students reported that the instruction had been sufficient. All programs could be completed within the time available.

The students reported that all programs were instructive and interesting. Most students reported that similar software should be applied more often. Before the introduction of balance and calculate, the professor applied training sessions to reach the same educational goals. During these training sessions the professor had to explain the same phenomena to each pair of students. On consequence of the introduction of the tutorials the time needed for this boring task decreased dramatically. The professor did not spend more than two hours instructing students during any of the sessions, even if four pairs of students were simultaneously using the programs. Consequently the professor preferred the application of the tutorials above that of a training session.

During the first field–test tempe was implemented in a purely open context. For many students this context was convenient. However, some students needed more instructions. During the second field–test, the program was used in combination with suggestions for simulated experiments, provided on paper. This context was convenient for all of the students.

Conclusions and discussion

The main conclusions drawn from our field–tests are:

Simple–structured, computer–controlled, text–based tutorials are suitable tools for teaching basic process engineering skills. The tutorials can be used stand–alone. The time needed for instructing the student decreases dramatically due to the replacement of sessions for training these skills by computer sessions. The professor and most students prefer computer sessions above training sessions.

Student–controlled, graphics–based simulations are suitable tools for teaching predictive modelling. The simulations must be used in combination with paper materials. The students and the professor like to use the simulations besides real experiments.

The construction of simple–structured, computer–controlled, text–based tutorials is cheap (The ratio between time needed for development and hours of produced courseware of the tested prototypes was about 10). These tutorials can be

implemented on cheap hardware (e.g., ibm xt compatible computers with a cga card). Because of these low costs, these tutorials are powerful tools for optimizing the teaching of basic process engineering skills.
Compared with the used tutorials, the construction of advanced simulations is expensive (ratio between time for development and hours of produced courseware was about 100). The simulations can be executed on the same hardware as the tutorials. However, in the field–test more advanced hardware was used (ibm compatible computers with an Intel 80386 microprocessor, a mathematical coprocessor and a vga card). This advanced hardware was required to obtain a convenient simulation speed and high quality graphics. For most simulations for process engineering education (at least) similar hardware will be required. This means that compared with the costs for creating and using tutorials, the costs for creating and using simulations are much higher. However, simulations can be used for reaching educational goals that cannot be realized by tutorials. Often alternatives for simulations are more costly than simulations (e.g., practical work in process engineering education often implies the use of costly installations and other costly resources). This means that simulations can be used in the process engineering curriculum for optimizing teaching and for the achievement of goals that cannot be reached by other means.

INSTRUCTIONAL TRANSACTION THEORY APPLIED TO COMPUTER SIMULATIONS

Willem Jaap Zwart, Department of Education, University of Twente, Enschede, The Netherlands

Computer simulation has great potential for use in education. However it's greatest advantage (freedom for the learner to interact with a model of reality) can turn into a disadvantage that makes it ineffective in terms of learning outcome. Students need to have a high prior knowledge to understand what they can do in the simulation environment, students might develop misconceptions (Burton & Brown, 1982) and the instructional environment isn't able to correct them, 'learning by doing' with simulations can be time-consuming and working with computer simulations can lead to fragmented knowledge of a system (Min, 1987).
The instructional transaction theory of Merrill (Merrill, 1991) provides an useful framework to decrease problems in the use of computer simulations. This theory describes the use of transaction shells to design interactive, not frame-based, courseware. It is based on the concept of a instructional transaction which is "the complete sequence of presentations and reactions to acquire a specific type of instructional goal" (Merrill, Li and Jones, 1991).
A transaction is an instructional algorithm. It uses three type of data:

- a knowledge base;
- a resource database;
- a set of instructional parameters.

Knowledge base
The knowledge base consists of frames. The instructional transaction theory describes three types of frames (entities, activities and processes) and a number of elaborations of these frames (attributes, components, abstractions and associations). Applied to computer simulations the knowledge base should include frames for the system as a whole, for every entity in the system and for the relations between these entities. Beside this an executable model of the system is necessary. Most of the time this is a mathematical representation of the process. The knowledge base can be called a conceptual model of the domain, the executable model a runnable model (van Joolingen & de Jong, 1991)

Resource database
The resource database contains mediated representations of the domain knowledge.

Specific for simulations the resource database holds techniques for *presentation, acceptation and communication* (Min, 1987). To use the knowledge in the (runnable) model for educational purposes the resource database contains presentation techniques to display the results of the model, acceptation techniques to make interventions in the model possible and communication techniques to transfer data from the model to the educational representation and back.

Instructional parameters

By configuring the instructional parameters of a transaction shell instruction can be designed for a particular population, learning task and environmental situation. The instructional parameters tune a simulation to the needs of specific learners, specific goals and learning strategy by defining

- the interventions,
- the output and the
- learning process.

The interventions can differ in number and type (e.g. choice out of alternatives vs. free, intervention only before simulation vs. changes during simulation) of interventions.

The representation of results can differ in which output-variables of the model to display and the way this representation is done.

The learning process can be free discovery learning, guided discovery learning, problem solving or scientific experimentation. Dependent on the chosen learning process the instructional designer needs to specify the guidance, the problems to present to the learner etc.

The value of the instructional transaction theory for computer simulations is twofold:

- the definition of domain knowledge to include in the simulation;
- the separation of the instructional parameters.

Ad 1

'Traditional' computer simulations focus on the relations between entities in a system. The representation of domain knowledge in a frame network makes clear that a mental model of a systems consist of more elements than the relations between entities.

Ad 2

The instructional parameters are usually set within the simulation itself, thus being an 'implicit' instruction strategy (Duffy & Jonassen, 1991). The separation of the instructional parameters makes this strategy 'explicit' thus offering possibilities to change it. This can lead to adaptive simulations (changes in the parameters during execution) and possibilities to re-use simulations in different educational settings (by changing the parameters in a preparation phase).

In the project "Instructional Design of an Optical DataBase (ODB)" at the University of Twente the Instructional Transaction Theory is applied.

The main objective of the ODB-project is to develop the didactics (instructional principles) for the use of multimedia databases for multiple educational targets. Within the Database dynamic objects (simulations) are incorporated. The

simulations can make use of the other parts of the DataBase like audio and video fragments, animations, questions and explanations.
The multiple use of simulations is encouraged by providing teachers the possibility to manipulate instructional parameters of the simulation. Doing this the teacher can adapt the simulation to his audience and his instructional goals.
From the simulation the learner can have access to other parts of the database including instruction in unknown concepts involved in the simulation and an encyclopedic system.
The current research focuses on the necessary elements to included in the domain knowledge representation and the identification and implementation of *instructional parameters* that influence the successful use of simulation in education.
For this purpose a 'process' transaction shell is being developed. Every educational computer simulation is a specific instance of this transaction shell based on a specific model, with specific in- and output possibilities and specific instructional events to facilitate learning.
By instantiating this class with specific instructional parameters in the context of a multimedia database it is expected that the instructional effectiveness of simulations can be improved.

Examples of Instructional parameters
Every transaction shell has it's own set of parameters. A set of parameters must be based on (empirical) research. In the work of Merrill parameters for the 'identify' transaction shell are given as an example (Merrill, 1991). The 'process' transaction shell has it's own set. The following parameters are examples of the parameters for this transaction shell. Described are the name of the parameter, the possible values and the way the influence the transaction.

Output mode
The output mode parameter describes the way results of the model are presented to the learner. Possible values are enactive, iconic or symbolic (Bruner, 1966).
As an example the output of a simulation of the renneting of milk is used:

Value	Output
Enactive	A sequence of video-segments showing a person checking the viscosity of the milk
Iconic	A graph of the viscosity of the milk against time
Symbolic	A column of the exact value of the viscosity at different moments in time

Example 1
Values of the output mode parameter
Order of the presentation

The 'order of the presentation' parameter describes the form in which relations are shown to the user. The values are zero-order, first order and quantitative (White & Frederiksen, 1989). As an example the amount of enzyme in milk during pasteurisation is used.

Value	Output
Zero	The presence or absence of the enzyme is shown
First	The change in the amount of enzyme is shown
Quantitative	The exact value of the amount of enzyme is shown

Example 2
Values of the 'order of the presentation' parameter.

Conclusion
The Instructional Transaction Theory can be applied to computer simulations. Research and development is necessary to indentify the instructional parameters and the domain knowlegde to include in a 'process' tranasction shell.

References

Bruner, J.S. (1966). Towards a theory of instruction. Cambridge, Mass: Belknap Press of Harvard University.

Burton, R.R., & Brown, J.S. (1982). An investigation of computer coaching for informal learning activities. In D. Sleeman & J.S. Brown (eds.), Intelligent Tutoring Systems. London: Academic Press.

Duffy, M, & D.H. Jonassen (1991). Constructivism: New Implications for Instructional Technology? Educational Technology, Vol , 7-12.

Joolingen, W.R. van, & de Jong, T. (1991). An instructional related domain representation for simulations. In T. de Jong & K. Tait: Towards formalising the components of an intelligent learning environment, DELTA project SAFE, delivarable SIM/19-20/A

Merrill, M.D. (1991). An introduction to Instructional Transaction Theory. In D.W. Dalton (Ed.): Proceedings of the 33rd Annual International Conference of the Association for the Development of Computer-Based Instruction Systems. Columbus (Ohio): Adcis.

Merrill, M.D. (1991). Instructional Transaction Shells: Functions, Methods and Parameters. Utah State University.

Min, F.B.M. (1987). Computersimulatie als leermiddel. Schoonhoven: Academic Service.

White, B.Y., & J.R. Frederiksen (1989). Causal models as intelligent learning environments for science and engineering education. Applied Artificial Intelligence, 3, 167-190.

A DYNAMICALLY CONFIGURABLE REMOTE INTERFACE FOR EDUCATIONAL COMPUTER CONFERENCING

D. Collins, & S J. Bostock, Department of Computer Science, Keele University, Staffordshire, United Kingdom

Introduction

"In any distance education system or service, communication between learners, and learners and tutors should be provided" (Opheikens, 1992) Students of distance learning or open learning courses need support from tutors and fellow students, and computer conferencing can provide this regardless of geographical separation or any requirement for timetabling, due to its asynchronous nature. While electronic mail generally involves private messages between specific individuals, computer conferencing supports, in addition, a variety of interactions between groups (See Harasim 1991 for a review).

Conferencing software simultaneously maintains a number of named conferences, possibly for different student groups. Each conference includes one or more topics of discussion to which individuals add their text messages, linking them to existing contributions. There is hierarchy of conferences, topics and messages, and each conference is managed by a moderator, such as a tutor. The conferences are maintained and are readable whenever a person joins the conferencing system which thus forms a resource for learning as well as a means of communication. Institutions of Higher Education in North America have used conferencing for ten years, and several in the United Kingdom have recently employed it, using for example the conferencing software CoSy and Caucus (Mason & Kaye, 1989, Hardy et al., 1991, Hartley et al. 1991). At Keele we have used conferencing for four years to support adult students on a variety of part-time courses in Information Technology including short courses and an MSc with students contacting the conferencing system from personal computers at home or at work (Collins & Atherton 1989). We are now extending its use to students on non-technical courses, such as MBAs, where students are even less tolerant of the difficulties of using conferencing . We are currently using the public domain NIX conferencing system developed by Lancaster University.

Difficulties with conferencing

We have identified two types of barriers to the success of the conferencing approach to teaching and learning: the technology and the interface.

The first problem faced by students is attempting to cable and control their modems correctly. We have not dictated to students the modems which they

should adopt for the course. Modems are still an expensive item in the UK and students with access to particular models should not be forced to buy new modems. Naturally, a related problem is that of the software which controls the modem and we have experimented with a wide range of communications software. Many problems produce inconsistency in the behaviour of communications software with specific modems. Most communications software is general purpose, and consequently has to be flexible in terms of configuration. In order to configure modems for use with a given communications service such as conferencing, users have to become familiar with an extensive range of new terms and concepts. This can be extremely frustrating for new users and is an obstacle to the real goal of using the conferencing system. Our students have found particular difficulty transferring files due to the need to run compatible software on host and remote computers simultaneously, and to select from a wide range of File Transfer Protocols.

Software user interface

A separate difficulty is caused by the combination of the potentially complex conference structure and a text-based interface, provided by the remote host computer while the local personal computer acts as a 'dumb terminal'.

The conference structures which can be created and traversed in contemporary conferencing systems are complex. Students often get lost in the structure and are unsure of the options available to them regarding navigation. Most conferencing systems have primitive navigational interfaces and only permit navigation between adjacent messages or topics in the branching conference structure. At low modem speeds, comprehensive help systems are often just as difficult to navigate and can be slow and therefore costly. Most navigational interfaces are designed as a compromise between clarity and verbosity. They tend to be either too terse to be useful for novices or too slow to be economic. Users are often presented with a stream of useless information which they are unable to stop because of output buffering in the host. Such problems are compounded by the student's knowledge that their telephone charges are mounting!

Interfaces to the host conferencing software generally fall into one of three categories: a command line with terse menu prompts; comprehensive text menus; or a bar-menu which responds to cursor keys to select menu options. These are considered in turn.

The command line approach based on terse menus is better suited to the expert user. Even for experts, however, the terse menu and navigational aids provide barely sufficient information to permit decisions to be made with confidence. Users are often constrained to operations at the current level of the conferencing structure without clear indication of a route to anything other than the next level up. Navigation is slow with repetition of menu prompts following each navigational operation.

Comprehensive text menus are common on personal computer applications but across conventional communication lines a typical text menu can take several seconds to generate on the screen. The slow speed greatly hampers more experienced users and, given the communications costs involved, should be considered too slow for all users.

Bar menus involve the use of screen/terminal control codes which require that host and remote computers agree on terminal emulation standards. We have students who have selected an inappropriate standard and have subsequently been unable to make any sense of host transmissions - nor are they consequently able to re-configure the host without re-starting the session. Stray, corrupt or missing control codes can have a devastating effect on the presentation, often resulting in the software failing, and resultant frustration. Error correction/detection software and hardware naturally improves this situation but generally at too great a cost for students. A separate set of menus has to be maintained on the host for each form of terminal emulation.

Friendly Interface to Conferencing System (FICS)

The types of interfaces discussed above have been developed in the mainframe culture of dumb terminals connected to the host by slow transmission lines. As all of our students use personal computers a much improved interface is possible. We have developed a communications and support front-end, FICS, to the host conferencing system NIX to improve the method of using a conferencing system. FICS runs on the local personal computer, connected to the host via a modem, and consequently its performance is not limited by transmission speeds. It addresses both technical and interface difficulties. It provides guidance to installing the modem, it automatically logs in to the conferencing software, it provides facilities to navigate the conference, does transparent file transfer, provides text editing for the preparation of messages, and it can be used off line with a local copy of the conference.

The principle problem involved in producing such a front end is the dynamic structure and content of conferences. FICS is automatically informed of the command structure, the message structure and the message contents when FICS is used to join the host conferencing system. It is thus a dynamically configurable remote front end to the conferencing software. For example, the addition of a new conference would simply require the transmission (transparent to the user) of the conference name and access rights. When the user chooses to join a conference, the changes in its structure since that user last joined are transmitted, to update the local 'map' of conference messages. An option to collect message contents means that the local map can be a full copy of the conference, which can be read off line and thus form a learning resource, especially valuable for new conference members. The addition of messages must be made live, on-line.

FICS uses the performance and graphic capabilities of personal computers, improving accessibility to conferences and allowing a more varied use of them for learning. The NIX host software required only small modifications, and can still be used in 'native' mode with a dumb terminal where FICS is not yet available and by new, unregistered users.

Future developments

With the support from the European Social Fund and Horizon projects, at Keele we have positively encouraged participation of disabled adults in all information technology courses (Billson et al. 1991). In addition to the general problems in using computers, the use of conferencing software has posed specific problems for disabled users. FICS will be modified to provide a variety of alternative interfaces

for different needs, for example, for blind users employing voice synthesis. The host software will not be aware of existence of these interfaces.
Another enhancement will be the facility to append graphics files to text messages. These might include diagrams or digitised photographs of authors. To display these requires a microcomputer and FICS will give an automatic display of images in the conference. It is also intended to produce a windows version of FICS.

References

Billson, A., Bostock, S. J., Collins, D., & Moore, S.A. (1991). Physically disabled students in information technology courses: problems and possibilities. Computers in Adult Education and Training. 2 (3) pp 206-212.

Collins, D., & Atherton, J. (1989). Remote learning with Bulletin Board Systems. Computers in Adult Education and Training. 1 (2) pp51-56.

Harasim, L.M. , ed. (1990). Online Education: perspectives on a new environment. Praeger, New York.

Hardy, G., Hodgson, V., McConnell, D., & Reynolds, M. (1991). Computer Mediated Communication for Management Training and Development. Centre for the Study of Management Learning, Lancaster University, UK.

Hartley, J., Tagg, A., Garber, B., Barry, D., & Fitter, M. (1991). Computer conferencing for distance education. Learning Technologies Unit of the Department of Employment, UK

Mason, R. & Kaye, A. (1989). Mindweave: communication, computers and distance education. Pergamon Press, Oxford.

Opheikens, B .(1992). Telecommunication in distance education: User demands and technological possibilities. 265-274 in 'Learning Technology in the European Communities', Proceedings of the Delta conference on research and development, The Hague, October 1990, Kluwer Academic Publishers, Dordrecht.

TELECAI: COMPUTER AIDED INSTRUCTION AT A DISTANCE

Jos van der Woude, CETIS, Hogeschool Utrecht, Utrecht, The Netherlands

Introduction to the problem

Higher education is confronted with a diminishing number of students. The Hogeschool Utrecht (part of the Dutch Higher Vocational Education system) meets this challenge by providing more flexible education: Computer Aided Instruction (CAI) delivered at a distance. Using network technology, an electronic link is setup among students and between students and the department. Our goal is to provide better and more flexible service for our students, the computer and network technology are the means to this end.

Method

The department of economics of the Hogeschool Utrecht provides part–time education for 2200 students. These students attend classes three evenings a week, frequently in addition to a full–time job. For a number of years now the department uses statistics courseware it developed itself. Students work with this material in the computer lab. This setup is not very flexible. The lab is constantly used to capacity, leading to a limited, and fixed, exposure time per student.
One way to improve flexibility is to provide computers to students at home. Students can then study the material at home as often and as long as they wish.
One crucial aspect of education, contact with the lecturer, threatens to disappear. We felt it important to continue this direct contact. Electronic mail links were set up with this purpose in mind.
We would like to stress the fact that electronic mail used in this way is a means, not an educational end. Our students are part–time business economics majors, not necessarily computer (and network) aficionados.
The TeleCAI concept was tested in a pilot study. A group of 20 part–time students received a computer with modem. Courseware for statistics was installed by the supplier, so these students could start immediately. Instead of the usual three evenings, these students only physically went to class two times per week, leading to a reduction in commuting as a side benefit. Furthermore an electronic mail facility was provided for tutoring and monitoring. Questions posed electronically to the lecturer were answered within 24 hours – including the weekend. All student action in the courseware was logged to a file. Upon completion of a session, these results were automatically uploaded to a central computer at the institute. The logs

of all students were, again automatically, inserted into a database. Once a week overviews from this database were produced for the benefit of the lecturer.

Results

The logistics of hardware/software/courseware/netware were all perfectly organized. We are convinced that we handled these aspects correctly. All material, except for the electronic mail, was extremely easy to use. There were some initial difficulties connecting the modems to the public telephone network. The problems ranged from having to dial through private exchanges to a outdated phone connection *without* a dialing tone. In one extreme case we paid for a new telephone connection.

Nearly all students (90%) were positive about this method of studying and wanted morecourses to be conducted in this way.

The students judged that they spend more time studying compared to conventional education. The role of the lecturer has changed significantly. In the classical system the effort of the lecturer is focussed on preparing and presenting the lecture. In the TeleCAI experiment the lecturer's activities were fragmented over the entire week, including weekends. This demands a lot of flexibility on the part of the lecturer. There is a clear need for further research into this. One of the more poignant results of the evaluation of this experiment is the apparent unsuitability of the electronic mail system used (MEMOCOM) for educational purposes. Students found this extremely difficult to use, especially in comparison with the rest of the software. Students found the lack of support for statistical and mathematical symbols in electronic mail difficult to accept. A final point of criticism concerned the fact that the electronic mail system has not been integrated into the courseware. Posing a question to the lecturer while working with the courseware involved the following string of activities: quit the courseware manager, load the electronic mail system, edit the message, send the message, quit electronic mail system, load courseware manager.

At the end of the year, in august 1991, 10 students had passed the exams and entered the second year. This drop–out rate of 50% was found to be comparable to similar groups of part–time students following courses in the classical way. We found a clear relationship between the amount of problems completed in the courseware and the final grades. Students who completed more than half of the problems in the courseware, all passed the exam.

Conclusion

Overall we are pleased with the results of the pilot. It is clear that this new way of studying demands a more active role from the student. We do not feel this is necessarily a negative aspect. The TeleCAI project has had direct spin–offs this year. All first–year students were offered a PC on a lease basis. To date 520 students made use of this offer. The result: currently over 600 students (including those who already owned a computer) work with the courseware at home. The computer labs in the institute will be phased out to a large extend. One aspect of the pilot, the electronic mail facility was found to be both technically and educationally immature and needs reworking. Electronic mail in its present form is too difficult to use for laypeople. For successful introduction in the educational process electronic mail needs to be as simple to operate as a lecturer and a book,

its main competitors. Of the two main functions of the electronic mail: automatic logging of results and tutoring, the former was the most successful.

Discussion

Presently we are developing our own electronic mail system in close cooperation with the people who implement the wide area network of the Hogeschool. It is to based on the X.400 '88 standard. All personnel and selected groups of students will have international electronic mail access. A helpdesk, to be set up shortly, will provide assistance with hardware (computer and modem) problems. We will continue to offer PCs, based on a four year lease contract, to all first year students. These PCs will come with approximately 30 Mbyte of courseware pre–installed for nearly all first year subjects. We will provide an automatic logging and analysis service for statistics and perhaps other subjects as well.

Management noted a potential difficulty with the scheduling of TeleCAI groups. Due to the modular nature of our educational programme students follow individual study routes. Students who follow a distance delivered course are hard to fit into the roster for the subjects taught classically. Because of the educational and practical difficulties with tele–tutoring, this aspect of the pilot will not go into the implementation phase next year as originally planned. Nevertheless, we will continue our research effort toward this end.Currently management is considering to offer our regular programme in the form of a correspondence course. The TeleCAI concept would fit here perfectly. Based on the experience gained in the TeleCAI pilot, we envision a 'study–station' with a graphical user interface. This machine will feature tools such as wordprocessors, spreadsheet packages, data–analysis tools and symbolic mathematics programs. These tools will integrate seamlessly with courseware and electronic mail facilities, making this machine an ideal study environment. While working with the courseware on such a machine, the student can cut and paste portions of the material and send it to the lecturer along with the message: 'I do not understand this'. If we were to name one crucial lesson learned from the TeleCAI project, it would be this: Installing more and more computers and computer laboratories in the institute is a dead end. After a critical point of computer use has been reached, it becomes simply more economical to provide each student with a truely 'personal' computer.

INCREASING COMPETENCE IN OPEN TELE-LEARNING ENVIRONMENTS

G. de Zeeuw, & H.W. Sligte, Center for Innovation and Co-operative Technology, University of Amsterdam, Amsterdam, The Netherlands

Abstract

Increasingly the need is felt to drastically change the structure of (secondary) education. This feeling usually is triggered by the fact that many societal processes for which education is relevant are speeding up, partly due to the use of new technologies, while education is lagging behind. This situation is interpreted as a breakdown in intentional interaction. It can be remedied by introducing structured computer mediated communications (SCMC), as part of a more general trend of increased use of CMC.

This paper sketches some results of CMC in secondary education, structured via 'teletrips', and performed under the aegis of the European Schools Project. CMC can be seen to increase the 'organizational openness' of education. It supports an increase of competence, in addition to an increase of performance.

Introduction to the problem

Idenburg, Nestor of Dutch educational policy research, in 1971, inspired by Husén, predicted radical changes in the system of education. He anticipated a growing flow of information. By what he called, 'branchings' of computers, it would be possible to access open learning centres from school, home or the factory. These 'branchings' now exist (Idenburg, 1971).

School education, as Idenburg envisioned it, would prepare pupils for life in a society that is constantly modernizing, imposing the need to speed up learning processes as well - that is by learning to find, by learning to learn, by learning to assimilate knowledge and practical competence, by learning to cope with increasing complexity. Such 'double' processes need adequate support to structure and coordinate the necessarily increased flow of information - both conceptually as well as technologically, and within and between social systems.

The authors of this paper have explored educational CMC within secondary education for the past three years, to support more dynamical forms of teaching, as well as the kind of structure needed to structure CMC. Teletrips have been found to provide an important core concept for such structuring.

Background

With new technologies, a major caveat is to avoid taking technology itself as a leading idea for improvement. Texts electronically transmitted and stored, infinitely revisable, capable of high speed transmission, and of being read by an indeterminate number of people at the same time, can make everyone our tele-conversation partners (Levinson, 1990). But to what avail? The main question remains to add to instruction, not to automate it.

To explore this additional element, we have noted,first, that education to a large extent has become *unsituated*. There is a breakdown in interaction, not only on the level of societal systems, but also on the level of the individual and his or her environment. Second, the distinction between *qualifications* and *skills* has become blurred. The concept of competence stresses the interactive interplay of increases of qualification and skill, precisely defined, in connecting parallel processes of action and development. People may be unable to learn a skill as long as they don't know the necessary qualifications; they may be unable to explore qualifications as long as they don't develop the implicit skills.

To support increases in competence, continued exchanges of information are necessary, both in terms of its content and of its structure. Skills have to be demonstrated; qualifications have to be reported upon. Such exchanges are called conversations, part of a collective endeavour. The concept of a teletrip is used to make such conversations possible. It is at this point that CMC may help to re-design traditional education environments.

The need for such re-design is taken to derive from the rule-based and 'culturally' determined superimposition of inflexible, closed organization principles and interaction patterns in schools. School is seen as uncapable to 'learn' at organizational level. In addition methods for inquiry and change based on traditional social-scientific epistemologies are perceived to be deficient, needing new approaches such as CMC (Van den Berg & Vandenberghe, 1988; Goodlad, 1984; Knoers, 1977; Lagerweij, 1986).

Method

Teletrips are co-operative projects, designed by partners from different areas and countries. Teletrips combine local learning and activities of enquiry, and conversations, both within and outside schools, using various types of media such as foreign languages, graphs, models and other behavioural instructive media. Under the aegis of the European Schools Project (ESP) over 50 teletrips have been performed over the past years, showing great variety over local conditions, conversation topics, and teachers' competence. ESP has evolved, through inquiry, development and organizational adjustments, from network, to infrastructure, into support system for those who want to explore this type of CMC.

Both at the level of a teletrip, and at the level of the support system, the approach aims to create and nurture grassroot growth. The interaction between the various 'actors' involved, i.e. (live ones) such as teachers, pupils, organizers, but also (functional ones) such as disciplines and topics, generates a 'larger world' for each actor - to test ideas, develop new qualifications and fields of study, and demonstrate the growth of competence (Bruner & Connolly, 1974).

Results

At the level of pupils an increased motivation for learning, better communicative language skills, enhanced intercultural consciousness, and sensitivity for the co-operative use of telematics, have been observed. Pupils learnt to deal with 'volatile' and 'silent' or tacit knowledge, knowledge not readily present in books or from teachers, especially involving frames of qualification to be agreed upon.
Results at the level of teachers involve improved interaction with pupils, integrating technology into one's teaching practice, increasing co-operation with colleagues, and intentionally designing and experimenting with unusual learning environments.
Improvements at school and inter-school level involve an increased level of awareness of the interactive advantages of different cultures, and of different ways of organizing learning.
These achievements are taken to signify the appearance of processes such as learning to learn, learning to find, learning to cope with social complexity. Students have been observed to learn to act both at the level of skills as well as on the level of qualifications.

Discussion

The results are still relatively modest, of course, compared to the aim of changing systems that are generally considered 'difficult-to-change'. Still, there are positive signs. There is resistance, for instance. Teachers indicate the demand of their colleagues, who 'inherit' the pupils, to stick to standard curriculum and to produce standard results. Clearly, new variation is being brought into the schools, challenging teachers to introduce co-operative ways of dealing with such variation, over and above the isolated learning processes that usually are considered necessary and sufficient. This development, in fact, realizes what Idenburg envisioned twenty years ago.
Indeed, a substantial change seems possible now, sometimes even referred to as a change of paradigm. Such a change would by necessity involve the design of new rules to observe what is needed for speeding up processes of learning, in the sense defined. Many of the present rules for high quality observation guarantee that, for example, learning to learn remains invisible. We can only 'see' learning, and only the learning of skills at that. It is in our schools that we find the most rigid forms of testing on visible skills and the most crude denials of the subtle forms of qualification that underlie competent action.
As we have become aware in recent times, such forms of qualification involve the use and development of implicit knowledge, the use of the fleeting thought and the various forms of mutual qualification that people so effectively use to organize and support their daily life. Fortunately, this type of looking at our environment is being explored increasingly in our societal organizations, be it computer supported or not. Coordination and interaction are turning into the mainstay of a more competent society (Kuutti & Bannon, 1991).
It is here that SCMC seems apt to fulfil a major role: it can make our schools and disciplines more organizationally open, on the alert for a combination of discipline and improvisation. Important areas for further study involve the structuring tools that may be used to elaborate and enhance competence, among them linguistic ones.

References

Berg, R. v. d., & Vandenberghe, R. (1988). Onderwijsvernieuwing op een keerpunt . Tilburg: Zwijsen.

Bruner, J. S., & Connolly, K. (1974). The growth of competence. London: Academic Press Developmental Sciences Trust,

Goodlad, J. I. (1983). A place called school. New York: McGraw-Hill.

Husén, T. (1979). The School in Question - A comparative Study of the School and its Future in Western Societies. Oxford: Oxford University Press.

Idenburg, P. J. (1971). Theorie van het onderwijsbeleid. Groningen: Wolters-Noordhoff.

Knoers, A. M. P. (1977). Algemene onderwijskunde voor het voortgezet onderwijs. Assen: Van Gorcum.

Kuutti, K., & Bannon, L. (1991). Some confusions at the Interface: Re-conceptualizing the "Interface" Problem. In M.I. Nurminen & G.R.S. Weir (eds.), Human Jobs and computer interfaces. Amsterdam: North-Holland

Lagerweij, N. A. J. (1986). Toekomst voor onderwijsvernieuwing. In A. Reints, & J. C. Voogt (Ed.), Naar beter Onderwijs (pp. 15). Tilburg: Zwijsen.

Levinson, P. (1990). Computer Conferencing in the Context of the Evolution of Media. In L. M. Harasim (Ed.), Online Education. Perspectives on a New Environment New York: Praeger.

STRUCTURED ELECTRONIC COMMUNICATION IN COMPUTER-MANAGED INSTRUCTION

Paul van Schaik, University of Twente, Enschede, The Netherlands

Abstract

One of the problems of Computer Supported Cooperative Work is the management of communication between computer users. Conversation structures in combination with communication tools for using these structures have been proposed as a solution to this problem. A disadvantage of these conversation structures is that they are fixed and consequently cannot be adapted to users' needs. As a remedy we propose adaptable conversation structures and adaptation tools for these structures. We describe the design and of a system for this purpose. The application of structured electronic communication that we consider is computer-managed instruction.

Introduction to the problem

In recent years there has been an increased interest in the possibilities that computers offer to support people who use computers in cooperative work. The software that supports this work is called groupware. Computer-supported cooperative work (CSCW) is (a) work that is suppported by groupware and (b) the scientific discipline that studies the development and use of groupware.

One of the problems of Computer Supported Cooperative Work (CSCW) is the management of communication between computer users. Conversation structures and communication tools for these structures have been proposed as a solution to this problem. Conversation structures formally define informal patterns of communication that people use in their work. A conversation structure for a particular class of communicative activity--for example communication between teacher and learner(s) in completing an assignment--includes a specification of roles, actions and their ordering, the sorts of messages to be composed and exchanged, and the organisation of the messages received. An example of a communication structure is the so-called 'conversation for action', in which one party makes a request to another, as described by Winograd (1988). The overall structure of this conversation is presented in the Appendix. For each point in a communication structure, the communication structure specifies a small set of possible action types, determined by the previous history of the communication. A communication structure manager programme, such as the Coordinator (Winograd, op. cit.), makes it possible to keep track of requests and commitments

545

made. In order to be useful communication structures must be tailorable to the needs of the group.
A disadvantage of systems that offer conversation structures is that these structures are fixed and cannot be adapted to users' needs.

A solution
As a remedy we propose adaptable conversation structures and adaptation tools. Conversation structures are adaptable when their constituting elements--such as roles, actions and their ordering--can be adapted. Adaptation tools can be used to manage conversation structures, for example adapting them or creating new structures.

Application to Computer-Managed Instruction
The application of electronic communication that we consider is Computer-Managed Instruction (CMI). In certain educational situations, such as in software programming education, a large part of the learner activities is performed using computers. In these situations one type of communication is between the teacher and learners about assignments, such as software programs. For example the course of the communication can be as follows: the teacher commissions an assignment, the learner commits himself to deliver the assignment, the learner delivers the assignment as completed, the teacher accepts or rejects the assignment as completed, etc.

The design of system for electronic communication
We distinguish two types of users of a software system for electronic communication based on the notion of adaptable conversation structures: users who communicate via conversation structures and users (communicating users) who manage conversation structures (managing users). We describe the design of such a system that we developed: task structure, data structure and the user interface.
The task structure consists of two main tasks: Converse (for communicating users) and Manage system (for managing users). For *Converse* there the following subtasks: Start a conversation, Continue a running conversation, View a directory of conversations, and View a directory of conversation types. For *Manage system* there the following subtasks: Adapt system to users' needs and Manage system access. For *Adapt system to users' needs* there are the following sub-subtasks: Create conversation structure, Adapt conversation structure, Create a variant of an existing conversation structure, Enable conversation structure, and Disable conversation structure.
The data structure consists of three parts. First there are *conversation structures*. Their purpose is to provide a common structure for related conversations. Therefore they represent the possible sequences of linguistic actions--such as "commission an assignment"--in a conversation, or, stated differently, "routes that the conversation can take". As an application for CMI a "Conversation for assignment processing" has especially been developed. The purpose of this structure is to support the process of commissioning, accepting, submitting, and grading assignments. Secondly there are *conversations*. Their purpose is to organise several related messages, for example for assignment processing.

Therefore they contain several messages that have been sent (and received). Each message corresponds with an action in a conversation structure. The structure determines possible next actions at each state in the conversation. Thirdly there are *messages*. By means of them the actual communication takes place. Therefore they consist of, among other things, a the name of the message sender, the message text, the identification of the message to which the message is a reply, and a pointer to the next message in the conversation.
The user interface largely reflects the task structure and makes it possible to switch quickly between (sub-sub)tasks.
Based on this design we developed a prototype system called "Compversations". Compversations runs on personal computers under the MS-DOS operating system.

Conclusions

We have shown that it is possible to design a system for structured electronic communication for CMI. Based on this design a system called "Compversations" has actually been developed. As an application, a conversation structure for assignment processing was realised with Compversations. Result from an experiment on the usability of the system will be reported elsewhere.

547

FORMAL FEATURES AS A DESIGN FACTOR OF VIDEO SEGMENTS IN INTERACTIVE VIDEO PROGRAMMES

P.W. Verhagen, Department of Education, University of Twente, Enschede, The Netherlands

Abstract

Video segments may be characterized by formal design features with respect to factors such as complexity of narration, mutual influence of picture and sound, use of super-imposed texts, information load due to technical terms. The presentation suggests ways to operationalize these factors and reports about an experiment in which the influence of these formal features was studied with respect to perceived information load of video segments by learners.

Introduction

Interactive video may be conceived as the combination of the interactive capabilities of the computer and the audiovisual power of video. This combination is generally considered as a powerful medium for delivering instruction, nowadays often incorporated in so-called multi media systems.

An essential property of adding computer control to video programmes is the possibility to stop for questions or excercises, or to give feedback to students about their progression, at any desired moment. A simple but fundamental question is: What is the optimum moment to stop for the purpose of achieving maximum learning? The literature shows some ideas with respect to this question (see for instance Bork, 1987; Laurillard, 1987; Schaffer and Hannafin, 1986). In any case, very short segments may limit typical video design factors to build a discourse to such an extent, that the communication possibilities of video are not used to its full potential. On the other hand very long segments may offer more information than learners can handle. Verhagen carried out a study to gather insight in this matter (Verhagen, 1992).

Preferred segment length

In that study, an experimental videodisc programme about cheesemaking was produced. This programme contains 252 information elements which form a connected discourse of 36 minutes if the programme is played linearly without stopping. An information element is defined as one uninterrupted statement of the narrator about which one factual question can be put. he controlling computer program is in a way, that the video programme can only be started or stopped

between two information elements. This causes video segments of the programme always to consist of whole numbers of information elements.
The programme was used to find out what segment lengths are preferred by learners. In the main condition of the study 117 subjects (university freshmen) participated. Each subject worked individually with the programme. This worked as follows:
The subject started the programme and then watched it untill he/she decided - for whatever reason - that it was time to ask for stopping by pushing a mouse button. The programme consequently stopped at the end of the information element in which the stopping was requested. The subject then answered all questions about the segment he/she just saw. Next feedback was given. For any question that was answered wrongly, the related video fragment was repeated followed by a second attempt to answer the question. As soon as feedback was completed, the subject was allowed to continue the programme from the beginning of the information element that followed on the one in which he/she stopped. In this way, the subject divided the video programme into segments by the repetition of watching, stopping, answering questions, getting feedback (with build in repetition of video parts with respect to missed questions) and starting the next segment.
The results showed self-chosen segment lengths of which the means per subject varied from 2.19 to 87.50, with a mean of mean length of 12.70 (standard deviation 11.77).

Segment length as a function of the position in the programme
Within the programme, the mean segment length appeared to vary as a function of the position in the programme: subjects as a group appeared to prefer shorter segments at one place in the programme than at other places. It could be that some parts of the programme are perceived as being easier than other parts. The presentation will try to analyse this phenomenon. It concerns the following question:
"To what extent influence formal features of video segments the information load of these segments as perceived by the learners?"
The formal features that are studied are:

- complexity of narration (with respect to the propositional structure of narrated sentences);
- the use of short supporting super imposed texts on screen (that were used in about one third of the information elements);
- the occurence of technical terms in the narration;
- the tuning of the narration to the images (sometimes not entirely correctly described as between channel redundancy).

In the presentation, the way in which these factors are rated and analysed are described and design issues will be discussed.

References

Bork, A. (1987). Lessons from computer-based learning. In: D. Laurillard (Ed). Interactive Media: Working Methods and Practical Applications. Chichester: Ellis Horwood.

Laurillard, D. (1987). Pedagogical design for interactive video. In: D. Laurillard (Ed). Interactive Media: Working Methods and Practical Applications. Chichester: Ellis Horwood.

Schaffer, L., & Hannafin, M.J. (1986). The effects of progressively enriched interactivity on learning from interactive video. Educational Communication and Technology Journal, 34, 89-96.

Verhagen, P.W. (1992). Preferred length of video segments in interactive video programmes. Paper presented at the 1992 AECT National Convention. (February 5-9), Washington, DC.

A VIDEODISC PROGRAM FOR THE ACQUISITION OF DIAGNOSTIC SKILLS IN HEALTH EDUCATION

J. Gulmans, & R. van den Berg, Department of Education, University of Twente, Enschede, The Netherlands

Introduction

In health education the main goal is to answer the question what kind of a disease does a patient have. The abilities which are prerequisite for answering this question are referred to as diagnostic skills. The central issue which will be discussed here is the design of a multi-media learning environment which will enhance the acquisition of these diagnostic skills. The assumption is that the design of such learning environment can be derived from prescriptions based on the *prototype* view of concept learning. This view will be elaborated firstly.

To provide a diagnosis is conceived as a type of categorisation. The process of categorisation refers to the subsumption of objects, processes and events in a category (Bourne, 1970; Bruner, Goodnow & Austin, 1956; Hunt, 1962). A category can have clear-cut boundaries or fuzzy boundaries. Most concepts in the medical domain can be characterized as *fuzzy* concepts (Zadeh, 1965). This can be illustrated by concepts as heart disease, low-back pain and headache. The question is what concept model is the most appropriate to represent the fuzziness of most medical concepts. Currently the two main concept models are the *classical model* (Bourne, 1965) and the *prototype model* (Rosch, 1973; Rosch & Mervis, 1975; Rosch, 1978). The prototype model seems the most appropriate for it takes into account explicitly the fuzziness of conceptual boundaries.

Tennyson & Cocchiarella (1986) have provided an instructional design model for learning concepts based on the prototype view of concept learning. They hypothesized that the learning of concepts consists of two phases:

- the formation of conceptual knowledge and
- the development of procedural knowledge.

From that background the instructional design guide-lines contain the following elements:

- a concept definition,
- a typical example,
- a set of expository instances,
- a set of interrogatory instances and
- a classification test.

551

Based on attribute characteristics (constant and variable dimensions) and the relational structure (successive and coordinate). Tennyson and Cocchiarella (1986) distinguished four strategies:

- constant plus successive,
- constant plus coordinate,
- variable plus successive and
- variable plus coordinate.

The videodisc program, presented in this paper, has been derived from the last strategy. It addresses learning coordinate concepts with variable dimensions. Especially it attends to the effect on learning outcomes of variation of instances and sequence of presentation of instances as well as the effect of contrast in variable attributes of instances.

One of the main characteristics of the instructional program is the difference between expository and interrogatory instances. The expository instances contain a linear video presentation with spoken comment. These instances refer to *typical* diseases.

The interrogatory instances encompass

- an initial video presentation where the patient is linear presented without the possibility for interaction,
- a secondary presentation encompasses the different aspects of the human body (E.g. face, hand, foot) as well as the monitor data (E.g. heart frequency, blood pressure, respiration, etc.).

In this phase the student can choose one aspect about which he will receive a question. Before answering this question he can optionally view the relevant video sequence.

Method

The reasons for choosing a videodisc program are:

- the possibility to use different information channels (text, speech and images),
- random access,
- branching,
- instant jump (Handbook Interactive Video, 1988).

The program*) contains three modules. In the first and second condition of the first module a best instance of the concept *shock* is presented followed by randomly presented interrogatory instances. Note that the first condition contains defining symptoms as well as irrelevant symptoms whereas the second condition contains only defining symptoms. In the third and fourth condition interrogatory instances are presented randomly without exposition of a typical instance. In the fourth condition the irrelevant symptoms are omitted. The hypothesis is that the presentation of a best example facilitates the formation of a mental image and that the enumeration of irrelevant symptoms will strengthen this effect. So condition one is expected to be superior.

The second module contains four conditions. In the first condition problems from one subcategory are practiced until no more errors are made. In the second condition a problem of each of the subcategories is presented followed by a problem of the next subcategory until all subcategories have been addressed to. The third and fourth condition are similar to respectively condition one and two

except for the fact that a relevant area of a variable is contrasted with a non-relevant area.
The third module attends to the problem of how new knowledge is related to existing knowledge. Relating new knowledge to existing knowledge enhances the retention of new knowledge. The assumption is that relating new knowledge can be accomplished by organizing a knowledge structure which can be presented before (first condition) and during practice (second condition) and at the moment that errors are made (third condition). The hypothesis is that the last condition is superior.
The videodisc program contains 30 patients ranging from typical to less typical instances of the three types of shock: cardial shock, distributive shock and hypovolemic shock. Each patient fragment is presented in a standardized way and contains an anamnesis, a list of more or less salient symptoms and monitor data.

Results
A formative evaluation has been completed and has let to various improvements as to screen design, content and presentation form of the content. A group of 10 experts in the field of eductional technology and with experience regarded to interactive video have completed a questionaire of 42 questions. The questions where subdivided in seven categories: information channels, sequencing of the instruction, questions, screen design, user interface, informational accuracy and attractivity. Among the most important improvements which were made are the reduction of the number of buttons, enhancing the interaction possibilities and tuning the information channels voice, image and written text.

Conclusion and discussion
It seems that a videodisc is a suitable tool for the development of the conceptual and procedural knowledge which are of crucial importance in the identification of diseases. Especially the interactive character of the videodisc as well as the possibility to confront subjects with clusters of symptoms seems promising for the health education as well as for instructional design for learning (medical) concepts.

*) The program was developed using the authoring language "Course of Action".

References

Bourne, L.E., Jr. (1965). Human conceptual behaviour. Boston: Allyn & Bacon.
Bourne, L.E., Jr. (1970). Knowing and using concepts. Psychological Review, 77, 546-556.
Bruner, J.S., Goodnow, J.J., & Austin, G.A. (1956). A study of thinking. New York: Science Editions.
Hunt, E.B. (1962). Concept learning: An information processing problem. New York: Wiley.
Interactive video: Introduction and handbook. (1988). London: The national interactive video centre.
Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In: T.E. Moore (Ed.): Cognitive development and the acquisition of language. New York: Academic Press.

553

Rosch, E., & Mervis, C.B. (1975). Family resemblances: Studies in the internal structure of categories. Cognitive psychology, 7, 573-605.
Rosch, E. (1978). Principles of categorisation. In: E. Rosch & B.B. Lloyd (Eds.), Cognition and categorisation. Hillsdale, NJ: Laurence Erlbaum Associates.
Tennyson, R.D., & Cocchiarella, M. (1986). An empirically based instructional design theory for teaching concepts. Review of Educational Research, 56, 40-71.
Zadeh, L.A. (1965). Fuzzy sets. Information and Control, 8, 338-353.

THE EFFICIENCY AND EFFECTIVENESS OF INTERACTIVE VIDEO IN SKILLS TRAINING: THE ENTERPRISE IN HIGHER EDUCATION EXPERIENCE

John Thompson, Centre for Executive Development, Ulster Business School; & Barry Commins, Department of Psychology, University of Ulster at Jordanstown, Newtownabbey, United Kingdom

Introduction

The Enterprise in Higher Education (EHE) initiative was started in the UK in 1988. The broad aim of the EHE (Training Agency, 1990) is that every person seeking a higher education qualification should be able to develop competencies and aptitudes relevant to enterprise.

Universities and Polytechnics were invited to apply for supportive funding of 1 million to be spent over a five year period. By September 1991 60 institutions had competed for and won contracts. The University of Ulster was awarded a contract in September 1990.

A survey of employers and graduates on the personal transferable skills required by graduates in the workplace (Harrison and Thompson, 1991) identified 16 skills, of which communication was the clear priority.

The University is therefore faced with the issue of how to develop personally transferable skills on 80 full time undergraduate programmes with 7,000 students over a five year period with over 700 academic staff. Although the University has a Social Skills Training Unit it has a limited capacity. The majority of staff have very limited or no experience in skills training. One option was Interactive Video (IV) using commercial courseware but evidence on cost effectiveness was required to obtain approval for expenditure. A survey of published studies revealed little evidence. There were few statistically evaluated studies and those were not in the area of 'enterprise' skills.

Two studies on the key skill communication (Thompson, 1992) indicated that IV was a cost–effective educational intervention (at the cognitive level) compared to lecturing. This paper reviews these results and presents a study which examines behaviourial as well as cognitive learning.

Analysis of Behaviourial Learning

Twenty six first year mechanical engineering students were randomly assigned to one of three interventions: interactive video (n = 10), lecture (n = 8), and control

555

with no educational intervention (n = 8). All subjects Ss completed three five minute videotaped two person role plays based on student/academic/ social issues. Intervening between the first and second role plays was either IV, lecture or for control group reading unrelated to social skills. The second role play occurred immediately after intervention and the third eight weeks after. The interactive video (IV) courseware (Face to Face Essential Interviewing Skills, IIS Ltd, 1987, London) focused on non–verbal communication, (eye contact, face, posture, head nods, tone and volume of voice, listening) and questioning (open, closed, leading).
Trained independent judges evaluated the appropriateness of the use of nonverbals and questions on five point scales. In order to control forexpectancy effects judges evaluated each individuals' performance without knowledge of intervention or role play order.

Results
Group analysis showed no statistically significant changes over time or intervention. Since a major objective of education and training is to develop skills in individuals, a secondary but descriptive analysis was conducted on a selection of case studies to ascertain the impact of interventions on individuals having different skill capabilities.

Case Study Profiles

	Interactive Video (IV)			Lecture (L)			Control (C)		
Performance	1	2	3	1	2	3	1	2	3
Good	36	43	35	37	42	31	40	36	32
Moderate	33	29	37	31	32	42	31	28	32
Poor	27	30	38	29	34	37	23	25	32

Table 1
Aggregated performances (max 50) of good, moderate and poor individuals in use of interactive video (IV), lecture (L), and control (C) conditions across three role play sessions, (1, 2, 3).

Cases were selected to represent good (GS), moderate (MS) and poor starters (PS) on the basis of initial performance (see Table 1). In both experimental conditions GSs improve between role play one and two with GSIV showing slightly better gains (7 vs 5). Subsequently GSs do worse, becoming the least effective with especially GSL displaying the greatest decrement. Both PSL and PSIV consistently improve across the three role plays with marginally greater gains being achieved by PSIV compared with PSL (11 vs 8). MSL consistently improves but MIV initially worsens but then improves significantly to slightly outperform MSL by the third role play. PSC consistently improves, MSC worsens then improves and GSC worsens throughout the learning period.
Analysis of profiles from first to third role plays across the ten skill areas shows that regardless of intervention, PSs improve in nonverbal skills (eye contact, facial expressions, posture and head nods) and for both IV and L interventions the use of closed questions. Only two skill areas show a loss, open questions (L) and

leading questions (C). MSL and MSIV achieve gains in the use of open, closed and leading questions, whilst MSC does worse on questioning. MSC also shows constant gains in facial expression, posture, head nods and listening. GSIV shows fewer losses and does show gains in listening, tone and volume of voice and open questions. GSC shows no gains and GSL shows only one gain, posture.
Overall nonverbal gains are achieved more easily whilst formal tuition may be necessary to improve in questioning. PSs respond consistently well even without formal tuition. MSs gain more consistently from the lecture and GSs lose out most without formal tuition but least under IV. It is clear that PSs have most room for improvement and do achieve this. Role play alone may function as both ateaching and motivational tool for PSs. GSs have less to gain and most to lose and this they do over the longer term. Perhaps GSs are initially well motivated and respond to formal tuition but later become complacent. They may be especially vulnerable in lecture only conditions.

Conclusions

The comparison of the initial two studies indicates that there are significant changes in cognitive learning. Although there were no statistically significant results for behaviourial learning, an examination of selected profiles reinforces the importance of considering the interaction of student capabilities with different learning environments.
There is a greater behaviourial improvement with IV for the weaker student. The comparison across students is a salutary reminder that the focus of EHE should be the individual student, not the total group. It is worthwhile surely if we can improve a student's performance by 40% on a one hour intervention by IV even if the changes are not statistically significant for the total group. There is the additional benefit that IV appears to promote longer term learning.
Although our costings show IV as cost effective, if the University is to invest in IV for Enterprise Skill Training there is a clear need to complete studies which involve a longer intervention than one hour and across a wider group of students. This is planned for 1992/93. For example, studies will examine (Binstead 1989) the use of role play with immediate student controlled video feedback as an alternative to IV and lecture and the role of the instructor (Wright and Dillon, 1990) in the promotion of learning.

References

Bosco, J. (1986). An Analysis of Evaluations of Interactive Video. Educational Technology, 26, 5.

Binstead, D. (1989). Developing Interpersonal Skills with Interactive Video. Training and Management Development Methods, 3, 1.

Harrison, R. T., & Thompson, J E. (1991). "The Content of the Enterprise Curriculum: Views of Employers and Graduates" – paper presented to the 2nd International Conference on Internationalising Higher Education, July, Dublin.

Norris, N., Davies, R., & Beattie, C. (1990). Evaluating New Technology: The Case of the Interactive Video in Schools' Programme. British Journal of Educational Technology, 21, 2.

Prosser, M.T. (1984). Towards more Effective Studies of Educational Media. Journal of Educational Technology, 15, 1.
Thompson, J E. (1992). "Teaching skills to thousands of Students" in Developing and Measuring Competencies: Aspects of Educational and Training Technology XXV, eds. Saunders D, and Race P, London: Kogan Page.
Training Agency (1990). EHE Report 1988 – 89, Sheffield: Training Agency.
Wright and Dillon. (1990). "Some Applications of Interactive Video in Initial Teacher Training", ETTI, 27, No. 1.

THE APPLICATION OF INTERACTIVE VIDEO TO THE PROFESSIONAL DEVELOPMENT OF TEACHERS ACROSS SCOTLAND

Martyn Roebuck, Scottish Office Education Department, United Kingdom

The paper is concerned with strategies for the use of multimedia technology in an economic and effective way for the continuing professional development of teachers. The Scottish Office Education Department over the past three years has commissioned the production of a range of such materials. These are designed to provide training in specific content areas, but at the same time to be a key component in the Department's approach to increasing the penetration of information technology into education. The paper describes the approaches adopted, accommodating issues such as cost, hardware availability, familiarity with technology, and access to quality staff development materials; and reviews the outcomes to date.

The main aim of the presentation is to indicate the rationale underlying the development of interactive video (IV) materials for teacher education in Scotland:

- to provide training in a stimulating way, in situ , for a significant proportion of Scotland's 46,000 teachers;
- to complement the other on-going Scotland-wide explorations of tutor-led, but largely print-based flexible delivery of staff development , by providing material at the front edge of multimedia applications;
- to involve teachers in the use of technologies which ultimately will have application in their classroom, and to give them some confidence in their use. (Note: the SOED policy has been to develop the use of IV primarily for in-service and pre-service teacher education rather than for classroom teaching - this is seen essentially as a later stage);
- to provide access to IV-based training on a large scale, by initiating the process through well designed materials, which teachers and headteachers would see as directly relevant to their own needs, and thus providing a hardware base for the use of other available materials and curriculum items;
- to explore and exploit the use of IV in a way which could be demonstrably cost-effective.

There will be reference to the criteria which underpin the design of the materials, such as those ensuring that:

- a wide variety of entry skills on the part of teachers or headteachers can be accommodated by each package;

- an individual user can use the materials in a number of ways, and on a cumulative basis on more than one occasion;
- the materials allow use by individuals, by groups, and in tutor-led contexts;
- the packages embrace a broad spectrum of relevant issues. (For example, "time management" may appear to be the main issue in one package, but it also addresses a range of management situations including the curriculum, personnel, racism. The teaching mathematics package may appear to be focusing primarily on delivering content and concepts, but it also provides a vehicle for analysing approaches to assessment and reporting individual pupil achievement);
- the hardware has the potential for a variety of other uses, including the delivery of staff development to teachers using commercial packages designed for industrial and commercial training; for the training of non-teaching staff; and for the support of management and information systems in schools.

The presentation will:

- set out and expand upon some of the objectives and design criteria for the packages, complemented by information in print form (briefly describing the relevant features of the packages, their form, use to date, and intended use, and cost);
- illustrate design features using extracts from an available package which simulates time management for the headteacher of a secondary school, and from one of three other packages under development and test for release during 1992 (on teaching mathematics at secondary level, and interviewing skills for staff appraisal in both primary secondary schools);
- conclude by evaluating progress in terms of the validity of the assumptions made, costs, customer reactions and criteria for assessing impact.

ARE THEY THINKING? PROBLEM SOLVING AND LEARNING THEORIES IN AN INTERACTIVE VIDEO APPLICATION

Madeleine J. Atkins, School of Education, University of Newcastle upon Tyne, United Kingdom

Introduction

This paper relates the findings of an empirical study of pupil use of an interactive video application to theories of learning. In particular, the paper will discuss the effectiveness of several design features of the disc which relate to behaviourist concepts of feedback and metacognitive theories of learning frameworks and learning skills.

The history of educational technology shows a strong behaviourist tradition through, among other things, the use of intrinsic and extrinsic feedback routines (Hannafin & Rieber 1989). Further, the nature of much computer programming has many parallels to the step by step, serialist approach of behaviourist learning theory. However, the cognitive, information - processing theories of the 1970s and 1980s have also now begun to make an impression on the design of multimedia applications by exploiting, for example, the use of surrogate travel techniques, object oriented programming and hypermedia database design. Rather than emphasising 'feedback', cognitive theorists have given more prominence to the idea of advice and metacognition. According to this approach, a principal aim of education should be to teach pupils how to learn, to develop their repertoire of learning skills and strategies and their ability to select from this repertoire appropriately for the task in hand. (Interestingly, there appears to have been little attempt to relate the concepts of feedback and metacognition theoretically.)

There is evidence from empirical studies of interactive video applications to show that, generally, feedback improves performance and the more elaborate the feedback the more learning is enhanced (Hannafin & Colamaio1987). However there is a trade-off for the designer to consider: elaborate feedback routines take longer for the user to complete than simple ones. Efficiencey of learning may decrease as effectiveness increases.

There is less firm empirical evidence on the learning efficacy of metacognitive frameworks in multimedia applications (Atkins 1992). Much appears to depend on the ability of the learner. If the learner already has well developed learning strategies then provision of a further metacognitive framework by the program can interfere with learning rather than enhance it. However, the reverse is true of low ability learners or learners whose repertoire of learning skills contains nothing

suitable for the task in hand. Without a firm cognitive framework such learners learn ineffectually from the application (Allred & Locatis 1988; Gay 1986).
The literature, while interesting, reveals one weakness. Very few of the interactive applications studied appear to have been designed to improve higher order intellectual skills or academic conceptual knowledge. Most have been designed for acquisition of factual information or for training in particular procedures. The videodisc evaluated in this study was therefore deliberately designed to improve pupils' reasoning and problem solving skills in relation to mathematical concepts. The disc employed both feedback and metacognitive features in its design.The question addressed in the evaluation was how effective these two approaches, from the different traditions of psychology, were in practice.

Method

The videodisc was designed for small group use by pupils without the teacher being present. It made use of visual contexts which were familiar or intriguing to young people. The pupils were involved in a narrative sequence which required them to sort out various problems. There were different types of problem on the disc : a closed 'code cracking' problem with one correct answer, a semistructured problem based on a business simulation with given parameters and several possible solutions and a more open problem centred on the difficulty of keeping buses running to time given different traffic conditions and route characteristics. Feedback was delivered in a variety of ways. On the closed problem it took the form of information (visual and aural) on the degree of correctness of the attempted answer. On the business simulation it came in direct comments on performance from the apparent managing director of the company, and indirectly from letters (read on screen) from satisfied and dissatisfied customers, from headlines seen in local newspapers, from conversations overheard on buses and even graffiti. The metacognitive framework was also supplied through a variety of means. Text instructions advised user groups to keep a record of their decision making as a basis for reviewing problem solving strategies, an on-screen clue facility was provided for the closed problem, and a filmed 'guide' of the same age as the users helped them to navigate through the events and options available.
The IV kit and disc was set up in an urban Middle School with 650 pupils on roll. The school was non-selective and was regarded as fairly typical in socioeconomic catchment area of many such English state schools. For this study six groups of four pupils were selected from the fourth year (12 - 13 year olds) with two boys and two girls in each group. The groups were selected to represent the widest range of ability possible on the basis of stanine scores on the Richmond mathematics test. The groups were withdrawn from their normal mathematics lessons for this study and were told they were assisting in the development of new learning materials. Each group used the disc on three occasions, each session lasting an average of 50 minutes.
The groups were videorecorded in every session. The videotapes were subjected to detailed visual analysis and transcripts of the pupils' conversations as they tackled the problems were studied separately. A category system of problem solving strategies was applied to the data and sequences of observed behaviour were coded accordingly.

Results and Discussion

The results of the analysis are interesting but complex to interpret (Atkins and Blissett 1989; Atkins and Blissett 1992). Motivation and mental effort were maintained at a high level across the ability range. Some refinement of problem solving strategy was evident. For example, there was a decreased reliance on unsystematic trial and error in the closed problem and a more sophisticated approach to the business simulation on the second and subsequent attempts. Nevertheless, it was clear that intrinsic feedback from the program was frequently ignored or used badly, with peer group interaction a more powerful factor than videodisc design. There was also some resistance to using the metacognitive framework provided and again, advice given to the pupils by the program was sometimes ignored particularly where the dynamic of the unfolding story was strong.

Paradoxically, some features of multimedia applications currrently seen as powerful motivators may, for some purposes, work against effective learning. The demand, for example, laid on learners to interact with the material, make decisions, see the consequences, and make further decisions may lead to a 'bias for action' which inhibits the taking of time to think through complex chains of reasoning or evaluate the success of a particular approach to problem solving. There was no intelligent tutor module incorporated into the application evaluated in this study. It is interesting to speculate whether such a module would have overcome some of the weaknesses identified. Similarly, the absence of a human teacher, while deliberately part of the design of this learning package, may have allowed suboptimal learning to go unchallenged.

In the presentation, these findings and paradoxes will be explored in more depth.

References

Atkins, M.J., Blissett ,G. (1989). 'Learning activities and interactive videodisc: an exploratory study'. British Journal of Educational Technology, 20,No.1.

Atkins, M.J., & Blissett,G. (1992). 'Interactive Video and Cognitive Problem Solving Skills : an investigation'. Educational Technology XXXII (1), 44-50.

Atkins, M.J. (1992). 'Theories of Learning and Multimedia Applications : An overview' . Research Papers in Education. Forthcoming.

Allred, K.F., & Locatis, C. (1988). 'Research, instructional design, and new technology'. Journal of Instructional Development ,11(1), 2-5.

Gay, G. (1986). 'Interaction of Learner Control and Prior Understanding in Computer-assisted Video Instruction'. Journal of Educational Psychology, 78 (3), 225-227.

Hannafin, M.J., & Colamaio, M.E. (1987). 'The Effects of Variations in Lesson Control and Practice on Learning from Interactive Video'. Educational Communication and Technology Journal, 35(4), 203-212.

Hannafin, M.J., & Rieber, L.P. (1989). 'Psychological foundations of instructional design for emerging computer-based technologies'. Educational Technology, Research and Development, 37(2), 91-101.

Reeves, T.C. (1986). 'Research and Evaluation models for the study of Interactive Video'. Journal of Computer-Based Instructio, 13(4), 102-106.

PROCEDURE FOR SUPPORTING THE SELECTION OF COURSEWARE BY SECONDARY SCHOOL TEACHERS

Baukje J. van Kesteren, Department of Agricultural Education, Wageningen Agricultural University, Wageningen; Gerrit J. Carleer, Betty A. Collis, & Tjeerd Plomp, Department of Education, University of Twente, Enschede, The Netherlands

Abstract

One of the main problems regarding the use of computers in education is the selection of courseware. For several reasons, courseware that may be eventually bought is difficult to access. In addition, there are difficulties relating to the support of the computer coordinator as the conduit through which preliminary information about courseware must flow as it arrives in the school. This study describes a dissemination procedure for a courseware selection instrument involving a workshop for computer coordinators and similar workshops (offered by the thus–trained coordinators) for their fellow–teachers. Results from six secondary schools and the implications of these results for the overall development of a framework and an instrument for courseware selection are described.

Introduction

Among the main obstacles affecting the use of computers in education are various difficulties relating to the process by which teachers select courseware. By 'courseware' is meant educational software accompanied by other materials such as work sheets for the students or guides for the teacher. Difficulties confronting the selection of courseware occur at various levels, including those relating to strategies for giving teachers a first exposure to subject–appropriate courseware as well as strategies for helping teachers make a more in–depth analysis. In addition, there are difficulties relating to the support of the computer coordinator as the conduit through which preliminary information about courseware must flow as it arrives at the school. Thus, there is an important need for strategies to provide teachers with better information about courseware and with some sort of aid for teachers to help them form a better perspective on the usefulness of different courseware alternatives. This report describes a component of on–going work in this occurring at the University of Twente in the Netherlands (Van Kesteren & Carleer, 1991).

Method

A theoretical framework based on the work of Fullan (1982) and of Fullan, Miles, and Anderson (1988) supported the selection of relevant variables affecting the effectiveness of courseware selection strategies and the possible relations between these variables.

Based on this analysis an inservice workshop for computer coordinators was developed as the central component of a dissemination procedure for better courseware selection. The computer coordinators were instructed in the use of an instrument for the selection of courseware by teachers called the 'Courseware Selector' and in strategies for the organization of a workshop on courseware selection for their fellow–teachers. The Courseware Selector consists of three parts, the first part guiding the search for existing courseware, the second part guiding the examination of the manual and (or) other information accompanying a given courseware package, and the third part guiding the examination of the courseware itself. In addition, information was given to the computer coordinators about different ways of organizing computer use in schools.

A multiple case study design was developed to compare developments in six secondary schools involved in this cascade approach. Each case included observations relating to teachers, computer coordinators, and school leaders. Attention was given to the process of courseware selection by teachers in each school at the start of the project, to the effect of the inservice workshop for the computer coordinators, to the impact of both of these on subsequent in–school workshops about courseware selection organized by the computer coordinators for their fellow–teachers, to the computer coordinators' activities during the course of the project, and to the sort of procedures used for courseware selection by teachers at the end of the study. Particular consideration was given to changes in the use of computers in the schools in general, and in the selection of courseware as it took place in the schools, in particular.

The data were mostly qualitative in nature, although within each case quantitative data were also gathered by way of questionnaires for the participating teachers. The sets of data were analyzed by seeking patterns in the material, much in the way Yin (1984) advises with respect to the formulation of 'case study questions'. Such questions were developed for this study, and related to the situation in the schools at the start of the study, the implementation of the use of the Courseware Selector, the effects of the project as a whole, and the conclusions that were to be drawn from the overall experience. Examples of these case study questions – concerning the situation at the start of the project and concerning the conclusions to be drawn from the overall experience, respectively – include the following:

- How can the use of computers in this school in general be characterized?
- On which information is this characterization based?
- By which further information is this conclusion supported?
- Which possible information is at variance with this conclusion?
- To what extent can one connect the positive or negative developments, or the absence of further developments during the project to the factors that have been described at the start of the project?
- To what extent is success or failure a consequence of the strategy for courseware selection that was offered?

- To what extent is success or failure a consequence of the way in which the computer coordinator put into practice the strategy for courseware selection?

At the completion of the study, the findings were linked to the theoretical framework.

Results

Although the effects of the overall strategy for improving the selection of courseware varied between schools, these effects were in general rather disappointing. Instead of enhancing the teachers' skills in selecting courseware, the main effect of the workshop that was offered to them by their computer coordinators was increasing their enthusiasm for the use of courseware. Although increased enthusiasm is not an undesirable effect in itself, it was not the effect that was most immediately sought. The main reason for an apparent lack of change in teachers' skill levels may be the fact that the use of computers in Dutch classrooms is generally occurring at such a low level that strategies to improve the selection of courseware do not have an adequate foundation. This seemed to be the case even in the six schools that were selected for the study (out of 80 that were willing to participate); schools that were comparatively far in computer–use experience based on national comparisons.

Conclusions

The main conclusions are as follows:

- The level of experience in the school relevant to computer use at the start of the implementation of a strategy for courseware selection is a critical variable in the success of subsequent implementation strategies. Schools or sections in which teachers have not yet used or selected courseware are not yet prepared for a strategy that focuses on the effective selection of courseware.
- The information about alternate strategies for the organization of the use of computers in schools that was given to the computer coordinators had different effects on the different coordinators. To those who felt relatively insecure about supporting computer use in their schools at the start of the project, this part of the course offered new insights about how to stimulate the use of computers by their fellow–teachers. In these cases, strengthening the functioning of the computer coordinator appeared to improve the results of the implementation of the overall strategy for courseware selection.
- In general, the effect of the overall strategy for the selection of courseware on teachers' intentions relative to courseware selection and use in the future was not very large. However, promising effects that did appear related to teachers getting better acquainted with courseware, increasing their enthusiasm for using courseware, and lessening their levels of 'computer–phobia'.

Discussion

The dissemination strategy for the use of the Courseware Selector that was used in this study is more effective in those cases in which teachers already have used and, preferably, have selected courseware before.

A workshop similar to the one which was offered in this study may be useful in schools which are less advanced in the use of computers, but the contents of this

workshop should relate to the use of computers in the classroom instead of the selection of courseware.

The decision to call in the computer coordinators to spread information and to teach skills regarding the selection of courseware appears to be a good one.

In the final report recommendations were given to schools regarding the organization of (aspects of) the use of computers in the school and in the classroom in general, and the organization of a workshop about courseware selection in particular. A few examples are as follows:

- Schools are recommended to organize a workshop about courseware selection only for sections in which courseware has been used already and in which some teachers already have selected courseware.
- The person who produces the workshop must be well instructed about the second and the third part of the Courseware Selector. The second part (the one guiding the examination of the manual) contains questions that cannot always be answered on the ground of the available documentation. This is important because it makes clear which information is lacking after reading the manual; however, if the participants are not informed about this possibility, this might provoke irritation.
- Workshops about courseware selection as well as workshops aiming at getting acquainted with courseware are best produced for complete sections. One reason is that consultation within a section plays an important role in the use of the Courseware Selector. A second reason is that joint activity in the field of computer use appears to have a stimulating effect on teachers.

References

Fullan, M. (1982). The meaning of educational change. Toronto: OISE.

Fullan, M., Miles, M.B., & Anderson, S.E. (1988). Strategies for implementing microcomputers in schools: The Ontario case. Toronto: Ministry of Education.

Van Kesteren, B.J., & Carleer, G.J. (1991). Een strategie voor courseware selectie door docenten. Enschede: Universiteit Twente, Onderzoekcentrum Toegepaste Onderwijskunde.

Yin, R.K. (1984). Case study research: Design and methods. Beverly Hills: Sage.

SUMMATIVE AND FORMATIVE EVALUATION OF EDUCATIONAL SOFTWARE IN EUROPEAN COUNTRIES

Leo W.F. Bonekamp, Educational Computing Consortium, Enschede, The Netherlands

Summary

Thorough evaluation is complex and expensive and much saving of effort could be achieved if Member States of the EC set up joint evaluation programms, or use the outcomes of each others evaluation programmes in their own decision making. The Community can help to ensure that a full range of appropriate evaluation tools is available to Member States and that evaluators are skilled in their use and in the interpretation of the outcomes. In order to enable such cooperation it is at first necessary to collect data on the current state of evaluation activities in Europe. This was the main goal of ECOSET.

ECOSET aimed at:

- an inventory of available evaluation instruments and procedures of application in the Member States;
- a comparison of these instruments and procedures with respect to the evaluation techniques, their cultural embeddedness and their curricular dependencies.

Evaluation is seen as 'the planned collection of information so that informed judgements can be made about the educational value of the microcomputer and the courseware used within the classroom' (Weston, 1987).

In the ECOSET–project data on evaluation activities of 19 organizations in ten member states of the EG are gathered by questionnaire. Key persons in four organisations were interviewed about evaluation. The participating member states were Germany, the United Kingdom, Belgium, Denmark, France, Luxembourg, the Netherlands, Portugal, Spain and Italy. Ten organizations do as well summative as formative evaluations and four organizations only perform summative evaluation activities. Five organizations are involved in formative evaluation activities only. From the data it is evident that national and local government have a significant role as commissioners in the evaluation activities of the organizations in the different member states. In almost all of the organizations (14) the formative evaluation activities deal with own development projects.

Fourteen organizations are involved in summative evaluation activities. At eleven the activities are commissioned by the government or by local authorities. The summative evaluation activities in the different member states can be placed on a continuum with on one side subjective, normative evaluation and on the other side the criterion–based approach (Dudley Marling, Owston, Searle, 1988). With a

'break–down' procedure we made a rough distinction in three groups of evaluation methodologies used in the European organizations, similar to three prototypical models of types of evaluation as proposed by Duchastel (1987). The first group are three organizations that do not use formal instruments for their evaluation activities. These organizations are more or less on the line of product reviews of educational software. A second group are thoseorganizations (5) that use formal instruments but the evaluation in real classroom situations is not a standard procedure. That does not mean that teachers are not involved in the evaluation activities. These organizations all use teachers as participants in the evaluation activities. This group probably has a more systematic way of evaluation with sets of mental categories and product characteristics with which they examine a software product. The third group of organizations (6) uses formal instruments and evaluates in real classroom situations as part of the standard evaluation procedure. The above mentioned division in evaluation activities does not mean that there are a lot of similarities in the organizations in one group. There can be for instance a lot of difference in the specifications of criteria used in the instruments, the number of teachers evaluating the program, the main function of the results etc.

Formative evaluation activities are standard at those 15 organizations that are engaged in software development. At two organizations the evaluators are independent subcontractors for parts of the evaluation activities in the decision phase and in the design & development phase. In the dissemination phase there is only one organization responsible for evaluation activities as independent subcontractor. In the decision phase there is one organization and in the design & development phase there are three organizations where the evaluators are as well member of the development team and independent subcontractor. Although preferable (Sparrow, 1973, Watson, 1987) it is not common in the organizations to have independant evaluators.

In the group of organizations performing formative evaluation activities two elements are remarkable. Organizations with few (<6) evaluation activities hardly train their evaluators. Almost all the other organizations give training to in house evaluators, subject matter specialists and/or teachers. If the number of evaluation activities gives an indication of the role of evaluation in the development process, you might say that those organizations also value training for evaluators as a necessity for good evaluations. The organizations with a lot of evaluation activities all use data of own or other observers. The aspect of training may also be connected with this fact. For observations the 'inter subjectivity' is an important factor. Inter subjectivity you will get by training your observers.

The data of the questionnaire do not give insight in the standardization of the different evaluation procedures in the development process. The interviews give an indication that formative evaluation activities differ per project and that procedures and methods of evaluation are planned and developed during the development process, instead of planned beforehand. Only one organization in the interviews indicate that the evaluation activities during the development process are part of a regularly applied quality assurance plan.

Conclusions and recommendations

Some conclusions of the ECOSET–project are:

- Future cooperation between the different countries of the EC have to concentrate on the optimization of the objectivity and the reliability of the evaluation methods and procedures. Although there are some parts in the procedure that have their cultural and curricular boundaries, there are other parts where cooperation is possible, for instance on the terrain of operational definitions of program characteristics, on test consistency and validity and on the development of procedures, tools and methodologies. The basis for cooperation should be a consensus about the target group and the purpose of the summative evaluation, the questions 'why?' and 'for whom?' (Smith and Keep, 1988).
- The underlying methods for evaluation in the Member States are relatively free of curricular and cultural influences. This implies that exchange of information between the Member States can take place in principle without to much difficulty in terms of'education–free' exchanges. From the experiences of the project it is recommended that the Commission stimulates setting up a course on evaluation methods to be distributed to interested Member States. Through such a course, unity in methodology will evolve, encouraging similar approaches towards evaluation resulting in compatibility of evaluation outcomes.
- In regard to the formative evaluation activities it seems logical that development projects and developing organizations will make use of the expertise and experiences of those organizations that perform a wide variety of formative evaluation activities in the different phases of the development process. One of the main conditions should be that the evaluation activities of those 'exemplary' organizations are formalized methods and procedures with good documentation, for instance in a 'quality assurance plan', of the different evaluation activities. It will facilitate the transfer of expertise to other EC countries.

References

Duchastel, P.C. (1987). Structures and methodologies for the evaluation of educational software. In: Studies in Educational Evaluation, 13, pp. 111 – 117.

Dudley–Marling C., Owston, R.G., Searle, D. (1988). A field–testing approach to software evaluation. In: Computers in the schools, 5(1/2), pp. 241 – 249.

Smith, D., & Keep , R. (1988). Eternal Triangulation: Case studies in the evaluation of educational software by classroom–based teacher groups. In: Comput. Educ. 12, No. 1, pp.151–156.

Sparrow, F.H. (1973). The role of the evaluator. In: Evaluation in Curriculum Development: Twelve case studies. London: MacMillan.

Watson, D. (1987). Developing CAL: Computers in the curriculum. London: Harper & Row, Publishers

Weston, P.R. (1987). Courseware development and evaluation. In: Beishuizen J. et al. The use of the microcomputer in teaching and learning. Swets & Zeitlinger b.v. Amsterdam.

A MODEL OF EVALUATING COURSEWARE

J.W. Wang, Department of Educational Information Technology, East China Normal University, Shanghai, China

Courseware engineering ' CE ', is directed at realizing ways and means for efficiently developing effective coursewares. For checking the effect of courseware, evaluation of courseware is one of its important step.

Courseware could be seen as products with three characters: instructional materials, instructional procedure, and computer' s application, so its evaluation would be a three-divisions measurement. If the measurement of materials which were checked out by editors is MM, and the measurement of computer' s techniques which could be tested in some normal testing procedures of software engineering is MC, and the measurement of the procedure is MP, then its evaluation should be:

$$E = MM * MC * MP \qquad (1)$$

For implementing this model, main problem is how to evaluate the procedure of a courseware. This paper is to try to find a model for evaluating the procedure of coursewares.

Hierarchical Model of Procedure

The model for evaluating the procedure of a courseware could be seen as a hierarchical network of several modules in different stages to do different analysis and collection of performance of students from a field-test.The lowest stage of procedure is unit stage which include a group of problems around certain topic, for example, some drills of skill in spelling.The next stage of procedure is lesson stage which is divided to several modules which are in unit stage and connect each others to form some instructional process so that the courseware could instruct each student a whole lesson through different passes of units.

Evaluating At Different Stage

Evaluating At Unit Stage

At unit stage, the main task of evaluation is collecting the data of student' s performance, making essential data processing, and doing basic analysis. If there are N problems in the unit and M students passed it, then problems and students could be listed non-ncreasively in line according to the orders of Pj or Si, where Pj is the number of students who did the Jth problem correctly and Si is the number of the problems being done by the Ith student. These data could be listed as a table called as the S-P TABLE in which Uij are the rightness of that the ith

571

student did the jth problem, and linking with the Si or Pj would form the S-LINE or P-LINE.

The most important data of S-P Table are difference of S-line and P-line (DSP) which discribes the difference between the abilities of students and hardness of the problems, its value is epual to the number between the S-line and the P-ine of S-P TABLE. It could be standardized as the difference coefficients DC by following form:

$$DC = = (DSP * 0.7) / (r * (l-r)) \quad (2)$$

in which r is the average right rating. Usually, DC is not bigger than 0.4, otherwise the problems should be checked again. And the Item Relational Structure Analysis could be done using the S-P TABLE. There are four data for each two problems Pi and Pj: a is the number of students who did both right, b is the number of students who did both wrong, b is the number of students who did Pi right and Pj wrong, c is the number of students who did Pi wrong and Pj right. The order coefficients Rij could be formed as following:

$$Rij = = l—(c * m) / ((c+d) * (a+c)) \quad (3)$$

Usually, if Rij>0.5, then we said that Pi→Pj, and it means that who could do Pj well must do Pi well, so the order of problems should be Pi pre Pj.

Meanwhile, another useful data are the correlative average time CTj and the correlative average rightness CRj, If the average time Tj and average rightness Rj of students for them are TT and TR, The correlative average time CTj and the correlative average rightness CRj could be get as follow:

$$CTj==DTj / TT; \qquad CRj==Rj / TR \quad (4)$$

Usually, the CRj / CTj would be increasing slowly when the courseware is suitable for this group of students. If CRj / CTj for problem j is too big or too small then these problem could be seen as too easy or too hard for the students. Normally,

$$1 / 2 < = CRj / CTj < = 2 \quad (5)$$

And if the CTj is decreasing quickly, then the order of problems in unit is not suitable for leaning of this group of students and should be fixed.

Evaluating At Lesson Stage

At lesson lesson stage, the task of evaluation is mainly to check out if its running could reach the object of this lesson.

At lesson stage, one part of evaluation is to trace the interaction between students and courseware to find out if there are any lone unit never being passed by any student in the group. Any solitary unit should be checked as it is not useful for students. Tracing the running process of a courseware could be done by recording all unit numbers which are passed by any student. If there are any unit number never be recorded, then this unit should be rechecked. And if there are any circle which repeat more than one time by same student then all units in it would be seen as too hark to be understood or not in right order.

Another part of evaluation at lesson stage is to measure the smoothness of total performance of the student group. It could be bone by comparing their record at each unit to see if there are any big bown / up jump between two units.

The Process Based on This Model

The process based on this model could be doing as following steps:

- Collecting Data: Let N students run this courseware, and collect their rightness and times for each problems as a matrix by a inserted module. Then the S-P TABLE would be made up by sorting the matrix.
- Checking the DSP of Each Unit: Using S-P TABLE Analysis procedure and form (2), the DSP and DC could be checked of each unit.
- Item Relation Structure Analysis: Using form (3), the Item Relation Structure Analysis could be done as wall, and it would be checked whether the order of problems in a unit is suitable or not.
- Time and Hardness Analysis: Using form (4) and (5), the tine and hardness analysis could be done as well.
- Tracing: List all unit numbers what have been run by at least one student in this group, so the solitary unit numbers and the repeatness of a circle of units would be recorded. And the smoothness of unit could be checked by comparing the scores of each students on two units.

Conclusion

This process based on this model could made the evaluation of coursewares easier and more objective. This is the first step for evaluating by field-testing, and we hope to develop some real auto-evaluation system as a part of CE in near future.

References

Zhang, J. (1988). Technology of Educational Evaluation. Liaoning Education Press.

Wang, J. (1990). The Model of Evaluating Courseware. Proceedings on the Fourth National CBE Conference in China. Tianjing, China.

Sakamoto. (Ed.). (1989). Analysis of Educational Data. Vol 3: Educational Information Science. Tokyo: First Copyright Press.

Fu, D. (1990). Computer and Education. Normal University Press: Centrel China.

FROM COMPUTERLITERACY TOWARDS INFORMATION SOCIETY LITERACY?

Jan Timmer, SCO, University of Amsterdam, Amsterdam, The Netherlands

Since the well known article of Andrew Molnar in 1978 and many other articles, the idea of computerliteracy became an important motive to introduce computers in the classroom. In the Netherlands the idea was implemented in 1984 in a five year plan called INSP, coordinated by the Ministry of education. However in practice computerliteracy became easily reduced to teaching students programming languages in stead of teaching them to master the interface between computer and user and even more important the interface between user and our changing informationsystems in society. It was predicted that those new technologies and new informationsystems would transform our society to an information society in which workpatterns, family life, leisure and entertainment and even the way we view ourselves as humans change. Alvin Toffler (1980) is one of the best known authors here.

At present the fear of a new form of illiteracy seems to have passed now. This does not imply a successful implementation of the INSP-plan. The process of introducing computers in classrooms (especially as a medium in the process of teaching), proceeds much slower than was anticipated. This is also the case with the introduction of personal computers in our society. In 1978 in the U.S.A. the subcommittee on domestic and international scientific planning, analysis and cooperation estimated a personal cost threshold for buying a computer between $ 200,- and $ 20,-. It was predicted that the first threshold would be reached in 1983 and the latter in 1988. Predictions are hard to make.

Some of the schools in the Netherlands that are leading the way in using the computer in education now tell that in contrast with to the 80 s, "computers" is are not used anymore to attract new students (Timmer 1991). This seems to be another sign that important motives to introduce computers in schools are not present anymore. As a consequence I could discuss more general topics:

- to deliver standardised knowledge to students;
- to support students in developing their own knowledge.

In respect to the second mode of learning, one can wonder wether we have to teach students more general knowledge about how to find and learn themselves the specific knowledge they need. In fact because knowledge in our society ages fast, education is asked to change from instruction to new ways of learning in which students become more active learners, asking themselves questions and searching for different resources and using those resources effectively in order to

answer those questions themselves. The computer can give a powerful support in that direction. This idea was not very much present in schools, though one school made an interesting remark about the problem of implementing new learning environments. In this school many discussions among teachers were reported about what structures one should provide to students. Too much structure confines the freedom of the students and hence very easily can reduce the competence of students. Too little structure however may lead to a situation in which students have to experiment too much. In that case students can easily "drown" instead of "learning to swim". It seems nearly impossible to solve that dilemma in a general way. Every teacher by "trial and error" has to find his or her own solutions to fit his own situation, values and criteria (his own competencies) as a teacher. Looking at the two modes of learning there is a continuous question what mixture of the two modes students should be provided with.

This conclusion can be reformulated by stating that structure has to be imposed in such a way that both the competencies of students and the competencies of teachers are increased. The implementation of new technologies in the learning environment of the student, adds new dimensions to this problem. One may also have to contribute to the collective competence of society in the future. This is the function that made education come to the fore in the first place.

Competence involves far more than skills. It includes both an activity and the reflection on this activity in a social group or organisation. According to de Zeeuw (Geijer and van der Zouwen 1986) the competence to perform a certain task is increased when new possibilities become available to impleement that task besides the old ones. Individual competence can be distinguished from collective competence which refers to the competence of a group of people or to the organisation or even society as a whole. In order to perform core tasks better and more efficiently, collective competence of organisations have to be increased. In many cases this problem is solved by increasing some individual competencies only. Unfortunately interactive processes in classrooms and schools then are reduced to "one way" activities which are imposed on students. The in-teractive process of teaching and learning (the second mode of learning) is reduced to instruction and the delivery of knowledge to students only (the first mode of learning). The teacher is not teaching anymore. Students are provided with structures that have to be used in exactly the same way as presented to them. There is little freedom for students to make connections with their own experiences and with other sources of information outside school. This must be deemed extremely important however in a society like in Western Europe and North America with environments that are very rich in information. A recent IEA-study (1989) told us that in these environments, in contrast to environments in the third world, the greater part of information is passed to the students outside schools. So it seems that the role of the school can be very easily marginalised in environments which are rich in information. It is my opinion that schools have an important role to coordinate learning outside schools with learning inside schools. In classroom interactive learning processes can be easily reduced to one way processes of the mode I type. The same holds for the interactive processes between education (schools) and society. At the conclusion of the INSP-programme in the Netherlands, Gevers (1988) pointed out that in the interaction between education and society, education should not only meet present needs of

society, but also contribute to the creation of future needs of society. This is like the brand new road that is build to reduce the traffic problems in a certain area. Because the road is present, it attracts traffic from other parts of the region as well and stimulates the inhabitants to buy more cars. So because of its very existance the road contributes to the creation of new problems and new needs. We have to realize that education has a responsability towards society in contributing to the creation of future useful needs and has a role in contributing to modifying and shaping our future society.

Our own ideas and knowledge shape the future. We know that in education, because of the "concrete here and now" the importance of ideas and underlying concepts that have a long term impact, and that often are not seen as concrete, is underesstimated. Ideas however are very important. In the battle in the pacific during the second world war, the Japanese navyleadership kept to the idea that battleships are the central part of a fleet. The Americans however took their experience from "Pearl Harbor" and from the sinking of the German battleship "Bismarck". They adopted the idea that aircraft carriers are the most important part of a fleet. They won the war in the pacific.

We know from the past (industrial innovation) that technological innovation in general gives rise to a redistribution of roles and functions in our society. Therefore profound questions arise about the future role of education and educational systems. Will schools marginalise or dissapear? If we look at schools as knowledge deliverysystems only, much can be taken over by distance education. If we look at schools as social organizations and as a focuspoint to coordinate learning (outside and inside school; mode I and mode II), it becomes a different story. A lot depends on the way we look at schools and the ideas we have about schools.

Predictions are hard to make. The future is unknowable. Therefore I put a questionmark after the title of this paper. What will become of education will be determined by what we are prepared to make of it ourselves. This again will depend on the ideas we have developed about education. Like in classroom structures have to be imposed so that both the competencies of education (schools and teachers) and the collective competence of society can be increased. Ideas are an important part of those structures. To stimulate interactive processes between education and society, ideas have to be developed, promoted, and discussed like:

- from passive learning to more active learning;
- school as a focuspoint to coordinate different kinds of learning;
- to find adequate mixtures of mode I and mode II learning.

There are signs that these ideas are not gaining weight in schools now. We still have to reach critical mass. Computers can serve as powerful supportsystems. After the motive of computerliteracy new motives have to be found via more general concepts and ideas.

CONDITIONS FOR EFFECTIVE IMPLEMENTATION OF COMPUTERS IN SECONDARY SCHOOLS IN THE NETHERLANDS: THE EXPERIMENTAL SCHOOLS PROJECT

Helma Vlas, Gerard Doornekamp, & Lidwien Cremers, Centre for Applied Research on Education (OCTO), University of Twente, Enschede, The Netherlands

Introduction

The lack of planning, coordination and workable strategies for the implementation of new technologies within secondary schools, contributes to the situation we can notice in the secondary schools in the Netherlands nowadays: ad hoc, isolated, individual efforts of a few enthusiastic teachers (about 10% to 15%) to use the computer in the classroom. Findings of a research project of the Centre for Applied Research in Education (University of Twente) in cooperation with two secondary schools indicate that much can be done to change the described gloomy situation. The identification of factors influencing the implementation of computers in the two experimental school and the design and implementation of strategies addressing those concerns has contributed to the fact that the use of computers in lessons in these schools, within three years, is almost triplicated (after 3 years 41% of the teachers uses the computer in their lessons), the use of a computer for word-processing (e.g. to create learning materials or tests) and the use of a computer for administration purposes by teachers has also increased by ample 300% (after three year the computer is used for these purposes by 51% of the teachers). Only 12% of the school staff are left who have not oriented yet on the use of the computer for educational purposes. In both schools the school administration system has been automated almost completely. Computers are often used by the clerical staff, deputy heads, the librarians and the career counsellors for both administration and management purposes. The way this research project is set up and some of the findings after the first three years are described in this paper. At the ECER'92 the most recent figures and a more comprehensive survey of the findings will be presented.

Research goals and method

The research project consists of two interrelated parts:

- To analyze, during four years, the circumstances of computer implementation within the two experimental schools, from an organizational, system theoretical point of view, within a framework of (theoretical) factors based on the accumulated

research on inno~vation implementation in education (in particular the ten critical factors introduced by Fullan et al., 1988) and to identify from this point of view the critical factors and critical implementation issues which apply especially for the Dutch secondary school setting. To plan, implement and evaluate, together with the planning team in the experimental schools, implementation support strategies addressing some of the critical factors.

- The data are gathered over a period of four years by means of yearly interviews with and written reports from those concerned and closely involved in the implementation process within the two experimental schools (like, among others, the computer coordinators, the system managers and the principals), analyses of the reports of the regular meetings of new computer consultative groups and analyses of the information provided by questionnaires filled in by all the teachers within the two experimental schools. The output will be twofold: 1) a scientific research report, scientific articles etc.; 2) a handbook for schools to guide the design of an effective implementation plan by providing better insight and a deeper understanding of the factors influencing the degree and quality of the implementation of computers in secondary schools and by describing possible solutions or problem solving strategies in a meaningful context to anticipate on specific problems and bottlenecks that can and probably will occur during the implementation process, based on the experiences within the two experimental schools during four years of implementation.

Some findings

Although the research goes on until 1993, some chief points that emerged from the findings can be notified already.

Principal's leadership and community support

In contradiction with the research findings of Fullan et al., principal's leadership and community support turned out to be of less importance in the Dutch situation. It is important that the principal facilitates and supports the implementation, communicates his concern, goals and priority, but he can share his authority and responsibility, for as far as computers are involved, with the computer coordinators as the principals in the two experimental schools did. The same thing applies for the schoolboard in the Netherlands. The board should be involved in the planning of the implementation process for reasons of support and facilitating, but they don't need to play any active role in the process. Community support turned out to be of hardly no importance for the success of the implementation. It's useful to keep the parents informed, but school don't have to involve them actively in the process.

Clarity and consensus about the change

It seems true, as Fullan et al. declare, that the success of the implementation will increase when implementators have gained a clear understanding of what to do in order to put the innovation into practice and when there is consensus within the school about the (importance of) the use of computers within the school. But experiences in the two experimental schools learned that for most of the teachers the understanding and consensus came only after they have had some experiences with the use of a computer, after they used the computer in the classroom and after they shared their experiences with their colleagues. And not

only a deeper understanding but also a tremendous enthusiasm. You need some fore-runners in the beginning, but not all the teachers have to be in the front line.

Quality and practicality of the innovation
Problems like: technical shorffalls, the lack of assistance for teachers, problems with the reservation of the computer labs are dealt with by introducing two new functions within the school: the system manager and the computer lab assistant. Workshops are organized to help teachers to select suitable courseware that fit in their curriculum. The computer managers for each subject department (a new function for one hour a week) select, demonstrate and introduce new courseware to be used by the teachers in their subject departments. Some problems still remain like the reservation of the computer lab (almost used to full capacity), lack of suitable courseware and time for teachers to practice and to experiment.

Professional development and assistance
Computer training within the school and demonstrations of computer applications for specific subjects (by colleagues) and the organization of continuing assistance (e.g. consultancy hours) turned out to be very important. In the experimental schools the training and assistance was organized and coordinated by the computer coordinators. The computer managers for each subject department organize demonstrations and support and stimulate the teachers in their own departments. Still, the enthusiasm among teachers has increased so tremendous that the schools by lack of time and budget, hardly can fulfil all the wishes for training

Environmental stability
In the Netherlands some major changes imposed by the government, like the decision to introduce a sort of comprehensive educational system within the existing structure and the introduction of a new financial system with more local responsibilities, have caused confusion and anxiety within the schools as a result of which the implementation of the computer has got less priority than needed. The government has until now neglected to conduct clear stated and unambiguous rules or policies to guide the implementation process or to provide a framework or target for direction. So schools, like our two experimental schools, have to set out their own goals, without having a clear picture of where the are heading for and which way to go. Not an easy accomplishment. There is hardly any extra budget or time allocation for the introduction of new technologies within the educational system. This imposes restraints on what can be done and lead to overload on those who voluntary take new tasks on their shoulders. Experiences learn that it is important that schools set bounds to their computer goals and decide on priorities and that they provide those actively involved in the implementation process with incentives like released time (when possible) or first choice by the distribution of new lessons.

Implementation monitoring and problem solving
It is very important to establish a strong structure for exchanging information, for decision making, for support and for coordination of all the activities and to make some people in the school responsible for the coordination and orchestration of all

the ongoing processes. The experimental schools did so by introducing the already mentioned new functions (especially the computer coordinators), new consultative groups, support networks and by arranging that computer issues become standard items on the agenda's of already existing meetings.

Conclusion and discussion

Surely, because of all the difference between school organizations, there can not be one "good" plan or strategy for implementation of computers that will work for all. Schools can only derive benefit from plans and strategies that meet their specific goals, wishes and requirements and are appropriate for their special situation. We expect that much more can be achieved by, based on the considerable knowledge that has accumulated from the research in the two experimental schools, providing the schools with deeper understanding and insight in the critical factors and implementation issues to be reckoned with and enable them to design their own operational plan for the implementation and integration of computers suitable to their own specific situation, by specifying design requirements and describing possible solutions for, or strategies to deal with, specific problems and bottlenecks and by describing the effects or impact to be expected of these solution or strategies within a meaningful context. The fore-runners schools, like our two experimental schools, have to learn by trial and error, but other schools can benefit from the considerable knowledge that has accumulated from the combination of research and practice. They can concentrate on all the important factors to reckon with right from the start and have more chances to succeed where others failed before.

References

Fullan, M., Miles, M. B., & Anderson, S. E. (1988). Strategies for implementing microcomputers in schools: the Ontario case. Toronto: Ministry of Education.

IMPLEMENTING UNIVERSITY CAL: A TOP–DOWN APPROACH

M.J.A. Mirande, & M.D. Leiblum, Instituut voor Onderwijskundige Dienstverlening (IOWO), Katholieke Universiteit Nijmegen, Nijmegen, The Netherlands

Introduction to the problem

The Computer–Aided–Learning (CAL) group, a section of the IOWO at the University of Nijmegen, has functioned for about 15 years as a central service agency providing assistance to all faculties on matters dealing with use of computers for instructional purposes (1). A reorganization several years ago led to policy changes in providing service, and was followed by the production of a major report requested by the university's governing board. This report (2) summarized the status of CAL at the university and made various recommendations including procedures that should be undertaken before project selection or usage decisions are made. A compact version of this report (7) is available in English. A chance to implement the new policies and procedures shortly followed with the start of a project (3) with the university's medical faculty. The paper discusses some procedures and tools developed, that can lead to a productive use of CAL reflected by a "top–down" approach to its implementation.

Methods

Four stages of CAL status may be identified, namely: unawareness, starting usage, explorative usage, and productive usage. Some characteristics of each stage are identified (6, p.6).

To reach the productive stage, an innovation strategy must be defined and adopted. A general strategy for implementing university CAL is specified and several management strategies and policies are also noted. The needs analysis and feasibility study play an important role in the CAL decision making process. The main theme of our implementation policy is based on the "rule of least resistance", which states that an innovation strategy has the best chance for success if it occurs in that part of an organization where implementation conditions are the most favorable.

The paper further discusses the "in–depth" implementation strategy for CAL usage within a total curriculum (in our case the medical curriculum). Five phases are noted: a) analysis of CAL possibilities, b) evaluation of existing materials, c) budgeting for new development, d) actual implementation, e) evaluation and reporting, along with sub–steps of each phase.

581

The presentation will focus on the first phase, e.g. Analysis. Criterion elements, for analysing CAL options, are discussed along with a related scoring system. Sixteen sub–elements are mentioned (Table 1), split among six criteria types: infrastructure, economic, educational, teacher related, student related, and discipline related. Supplementing this is a categorization of expected qualitative and quantitative benefits divided among processes and results.
The first analysis performed was a university review of current CAL status, per faculty (2). The Medical Faculty scored highest for CAL "potential" among all faculties. The "Global Curriculum Analysis" (3) was the second study (Table 1; elements 5–8). It reviewed data on numbers of students, stability of course contents, final examination results (pass–fail ratios) and course evaluations. Based on analysis, the first two years of the medical curriculum seemed most appropriate for CAL applications. The "Detailed Curriculum Analysis" (4) and "Specific Analysis of course units" followed. It was decided that each course or module should be evaluated for possible CAL applications (Table 1; criteria 15–18). For the latter, teaching staff received questionnaires and many were personally interviewed. The goal here was to specifically select those study units, e.g. teaching modules, laboratories, applied medical training, "free–choice" specialities, and at a more advanced level, "co–assistantsships", that "qualified" or had the most chance for CAL success (e.g. based on study of available literature, questionnaires, and interviews). All of the 18 selection criteria were applied. Per criteria element, questions and rating scales were designed to comprise the total printed form. Each returned questionnaire was scored and added to interview reviews where available. Those selected came under consideration for yet another more detailed interview or for inclusion in a later to be carried out, "courseware search" (5). We considered the teacher's attitude towards CAL, its utility, and his willingness to cooperate in later joint efforts to be of extreme importance. Table 1 presents an overview of selection criteria distributed across four types ofanalysis.

Results

The methods previously described were put to use at two higher educational institutions (one university and one vocational school) and current requests include (1993–94) application in 3 other disciplines.
One hundred and twenty eight (128) questionnaires were sent to medical study coordinators. After reminders, 77 were returned and processed. Almost one–third (26) of the medical disiplines (represented by their coordinators) came under consideration for further study. At least one–fifth of the courses or modules seemed ripe for further analysis.
From second interviews with the 26 staff members it was determined that 12 disciplines had real "interest" and satisfied selection criteria for "original" (newly developed) CAL applications. Teachers involved in the remaining disciplines were primarily interested in finding "ready–made" courseware, mostly because they feared that original CAL development would cost too much of their time.
The most common motives for using CAL fell under four categories: improved training & preparation for students (to do the job better or provide remediation), more frequent practice sessions, more effective problem solving for specific subject matters, and time savings for teaching staff.

The twelve (sub) disciplines were selected based on economical, educational, and teacher related criteria. In other words, they were (sub) disciplines where, over a period of five years: relatively large student numbers were expected, teachers were willing to invest development time, and where the staff felt that the "payoff" would be significant. Based on these results, an implementation plan (over several years) has been established, original courseware produced, and courseware searches have been carried out. While final evaluation is not yet completed, first indications are that that the used methods and procedures are effective.

Conclusion

To make an impact in higher education, CAL usage must reach the "productive" stage as described in this report. The methods and tools used to implement this stage have been tried and tested in actual projects. Indications are that with appropriate internal support, other training or education organizations may use similar approaches, leading to a more efficient utilization of CAL in higher education.

Discussion

Growth in the use of CAL has been slow but not spectactular in Dutch higher education. The old saying about being able to, " take a horse to water but not make it drink " is applicable. At the primary and secondary educational levels, some national (government) support has been granted, but higher education has not benefited from major subsidies. While earlier individual initiatives (by teachers themselves) have been many, few "top–down" initiatives, via departmental heads, educational committees, or governing boards have been realized. At Nijmegen, this approach was made possible by a three year "CAL innovation" project supported by all of the previous entities. Now in its second year, interim findings give us reason to believe that the "productive" stage of CAL will be successfully reached.

References

Leiblum, M.D., Hermans, D., & Derks, K. (1986), " A Decade of CAL at a Dutch University ", Computers & Education, 10, No. 1, 229–243. Pergamon Press, U.K.

Leiblum, M.D., & Mirande, M.J.A. (1990). Computerondersteund onderwijs aan de KUN: Aanbevelingen, stand van zaken, achtergronden (CAL at KUN: Recommendations, status, background. Report), IOWO, KUN.

Mirande, M. (1991). Globale Curriculumanalyse: 1e voortgangsverslag van het innovatieproject, implementatie COO in de faculteit de Medische Wetenschappen (Global Curriculum Analysis: 1st Project Report of the CAL Medical Faculty Innovation Project), IOWO, KUN.

Mirande, M., & Leiblum, M. (1991). Gedetaillerde curriculumanalyse: 2e voortgangsverslag van het innovatieproject, implementatie COO in de faculteit der Medische Wetenschappen (Detailed Curriculum Analysis: 2nd Project report), IOWO, KUN.

Leiblum, M.(1991). Courseware Search: 3rd Progress Report, IOWO, KUN.

Mirande, M., Leiblum, M., & Wilmsen, M. (1992). De computer en het curriculum: Naar produktief gebruik van COO (The Computer and the Curriculum: Towards productive usage of CAL), IOWO report and brochure, KUN. (Dutch, available on request).

Leiblum, M. (1992). Implementing CAL at a University, in: Computers & Education, 18, No.1–3, 109–118, Pergamon Press.

Appendix

Table 1

Organisational units and selection criteria elements for analysing the implementation of CAL.
(Translated from Leiblum (1992), p. 19)

Criteria for Applying CAL	• • U&N: Res. across Educational Institution	Global Analysis of Faculties	Detailed Analysis of curriculum	Analysis courses-modules
Infra-structural Criteria				
1. Current amount of CAL usage	•			
2. Computer facilities	•			
3. Existing CAL policies	•			
4. Financial resources	•			
Economic Criteria				
5. Student numbers		•	•	
6. Content stability		•	•	
7. Cooperation possibilities				•
8. Subsidy possibilities				•
Educational Criteria				
9. Type of instruct. Goals and Tasks	•		•	•
10. Openness to educ. improvements		•	•	•
Teacher related criteria				
11. Acceptance by teacher			•	
12. Committment of teacher			•	
Student related criteria				
13. Usefullness to student				•
14. Student accessability				•
Discipline related criteria				
15. Representability: clinical–non clinical		•		•
16. Importance of visuals in non clinical courses				•
17. Clinical reasoning in clinical courses				•
18. Relevancy to basic courses				•

• • U&N = User & Needs Analysis

COST-EFFECTIVENESS ANALYSIS OF THREE ALTERNATIVE WAYS OF MEDIA USE TO PROVIDE ELEMENTARY COMPUTER USE TRAINING

Joachim Wetterling, University of Twente, Enschede, The Netherlands

Introduction

The personal computer can be found these days in almost every office and also in the houses of many families. Because of this there is a great demand for opportunities to learn to work efficiently with computers. But there are cost and effectiveness differences between the many alternatives of delivery of instruction that are available for the learning of personal computer use. These are related to student characteristics and environmental and instructional aspects, that can influence the decision of which method of training delivery will be used. Growing complexity in this area of decision making leads to a need for more assistance.

Problem

There is a need for more and better cost-effectiveness analysis before choices are made for application of interactive media in education and training. These decisions can be based on economic as well as educational motives. Many decisions regarding educational media application however are taken without a explicit cost-effectiveness analysis. Instruments may be used in cost-effectiveness analysis that assist the decision maker. Instrumentation can also lead to a growing use of cost-effectiveness analysis and as a result to a better justified media use. Functions an instrument can assist the decision maker in are: (a) letting him compare cost and effectiveness aspects of different alternatives, (b) performing more complex calculations, (c) transferring combinations of data to decision proposals, or (d) just to make him think more clearly about the analysis he is doing can have a great value. To do this the instrument must be both flexible and accurate to avoid mismanagement as a consequence of wrong interpretation of information.

In this paper three ways to provide elementary computer use are compared on their cost-effectiveness. Implications of instrument use for decision making in each of them will be given. Choices of institutions are reviewed that have been made for one of the three alternatives. Choice influencing aspects are costs for equipment, facilities and instructors, target group (social, intellectual and physical level), instructor experience and training content level. Also the commercial situation of the institution that provides the training compared to other similar institutions has

influence on the media choice. There is less attention to the effectiveness, probably because it is so difficult to define, let alone measure the effectiveness of the use of specific media in education and training. Clark and Sugrue state that it does not matter what media you use. It is much more important to have a sound instructional strategy where the media are implemented so that they serve to the goals that you want to reach in the instruction (Clark & Sugrue, 1990). Offir and Katz think that the way learners learn to work with media used for instruction is important for media choice decisions connected with available time, resources and learning goals (Offir & Katz, 1990). Next to this there is more attention lately on intangible effects directly and indirectly related to media choice and use.

Method

The topic of this paper is cost-effectiveness analysis of different methods of learning to work with personal computers as an aid in decision making. In costing educational media use often the ingredients method is used.
Levin gives a thorough introduction in this costing strategy (Levin, 1983). With regard to effectiveness, prior knowledge, instructional and learning strategies, demands by the society or the employer and motivation play an important role in this area of training. The paper will use dutch projects as illustrations for this discussion. In the presentation cost-effectiveness issues regarding the three illustrations will be discussed.
Different ways of evaluating elementary computer use training will be given and the value of them with regard to different delivery modes and instructional strategies.
The first alternative is the use of computer based training. People can learn with this technique at a place or time that fits them most. This way of training is characterized by high development costs and a relatively short lifetime (Crombag et al.. 1979). Also learners need computer equipment and facilities. However, it is also characterized by low operating costs and when implementation is optimized very little time losses due to training have to be suffered. An example of computer-based training in elementary computer use is a course developed by Courseware Midden Nederland that is mainly sold to companies for on-the-job use. (Courseware Midden Nederland, 1991). Main advantages of computer-based training as delivery mode are the flexibility and interactivity for the learner. Companies can decide to choose for computer-based training when a large amount of staff has to be trained that already use a computer in their job and the training can be easily embedded in the job. Companies save travel and accommodation costs next to production losses when staff has to leave their job for the training.
The second alternative is the use of mass media (television, radio) for the delivery of courses. People watch television programs and listen to radio programs in addition to course material they work with at their own computer. Although less than in computer based training, there is a certain amount of freedom of time and place to learn left to the students. High development costs are typical for this delivery mode but also the costs of letting the instruction come to the student are higher than in traditional instruction environments. Because of the large amount of learners that use television courses the costs for the learner are low. Instructor related costs are very low once the course is in operation but equipment costs are

high for the learners too for they need at least a television set. A personal computer is needed for the practical exercises. An example is the course "PC-Prive" that Teleac, a dutch institution for distance education using television and radio, has offered (Teleac, 1991). Everyone can attend this course. Main advantages of television as delivery channel for elementary computer use is that large groups that otherwise are not able or willing to learn to use computers now can be reached and the relatively low costs for the learner. The main target group of this alternative is the individual at home for instance housewifes, unemployed or people who don't use a computer in their job but want to learn to work with it. The main reason for choosing television courses for learning to work with the computer is that enrollment is cheap, you do not have to leave the house and you can study on your own pace.
The third alternative is the use of classrooms where people have lessons in groups given by a teacher. In these classes every trainee can use a separate computer to experience working with it and for doing exercises. Students must come to these classes on scheduled times. Cost accents in this alternative are the personal costs of the teachers that provide the instruction. Learners or their employers have the fees as costs and sometimes travel and accommodation for the learners. Examples of this delivery channel are the many institutions that vary in size and are spread over the whole country. Often the fee for attending a course here is high so many learners in this alternative are people that have been sent here by their company.
But there are courses too that use this strategy for example for unemployed who want to improve their value on the labour market by learning to work with computers. In this case the courses are sponsored by the community. An example of an institution that provides training in elementary computer use by means of instructor-led training in classroom situations is Micro Teach Opleidingen (Micro Teach, 1991). Main advantages in this alternative are the individual attention the instructor can give to specific problems of the learners. Also fast and relevant feedback is possible here. Companies can choose for this delivery mode when they think the personal feedback and the specific instructional environment to be important. For relatively small amounts of learners this alternative is probably less expensive than the use of computer-based training.

Discussion
Analysis of these three methods has shown that regarding final cost-effectiveness analysis results there is not much difference between the alternatives. The alternative private users choose will often depend on personal motives. In corporate training often the method will be chosen that fits most to the regular job of the trainees. The hypothesis is that also cost aspects are more important for private users while effectiveness aspects are more important for business users except when very excessive sums of money are concerned. Cost-effectiveness analysis of media use in education and training is affected by a complex combination of instructional strategies used to reach the goals, economic constraints for the training, environmental aspects that influence the media choice in an educational situation and personal aspects of students and instructors. Decision makers have to be convinced that this rational analysis has to be done

587

before making decisions about application of media in education and training and that it has both an educational and an economic value.

References

Clark, R.E., & Sugrue, B.M. (1990). North American dispute about research on learning from media. International Journal of Educational Research, 14(6), 507-519.

Courseware Midden Nederland. (1991). Praktisch Computer Gebruik. COO Cursus. Utrecht.

Crombag, H.F.M., Chang, T.M., Drift van der, K.D.J.M., & Moonen, J.C.M.M. (1979). Onderwijsmiddelen van de Open Universiteit: functies en kosten. Den Haag: Staatsuitgeverij

Levin, H.M. (1983). Cost-Effectiveness: a primer. Beverly Hills: Sage.

Micro Teach Opleidingen. (1991). Studiegids 1992. Enschede.

Offir, B., & Katz, Y.J. (1990). The learning curve model for analysing the cost-effectiveness of a training system. Education & Computing, 6, 161-164.

Teleac. (1991). Cursusmateriaal van de cursus PC-Prive. Utrecht.

THE PORTABILITY OF EWB PACKAGE TO CHINA

Z-P. Dong, & J-P. Zhang, Department of Education Information Technology, East China Normal University, Shanghai, China

The EWB (Electronics Workbench) package was developed by the Interactive Image Technologies (IIT) Ltd. of Canada. In order to explore the portability of EWB package using in China, the IIT offered us a EWB copy last year. After running this package for several months, we feel it is an ideal environment and tool for learning analog and digital logic circuits. We also find this package is flexible enough to keeb pace with most textbooks and courses in China. It is much faster for teacher to make demonstration materials and asves time to prepare circuit diagrams of tested circuits for exercises and experiments. Of course, if the package is truly possible to be acceptable and have a good market in China, some adaptations and investigation should be done. So we have made a preliminary plan for this project:

- A simple computer system in classroom has be build in which allocated a pc computer and several monitors for running EWB during the courses of analog and digital circuits. A set demonstrations of experiment and exercise for courses are being developed by teachers.
- Organize the students to use the EWB after class. the students will use the EWB to review the course, to do exercise and design circuit.
- A special teacher group for investigation is necessary. They come from the Department of Education Information Technology, Electronics Science, Computer Science and Vocational Education School.
- Hold one or two EWB demonstrations to national CAI and CBE conference, for getting all participants responses and opinions.
- We made a questionnaire of the portability of EWB package to all users, then the responses and opinions will be collected in detail.

The questionnaire as follows:

The portability of EWB package to China questionnaire

Name:
Professional:
Special field of study:
Date:
A. Please choose (use" x ") according to your opinions:

Item \ Respons	A	B	C	D
Operate and use very convenient				
To learn and design circuits very convenient				
The result as the preparatory object				
No mistakes on running				
Main experiments of electronics can be finished				
Save time for circuit design and experiment				

B. Please write down your opinions as follows items:

(I) Use value:

(2) Problems:

(3) Do you want to buy this package if the problem above can be solved:

(4) The price range is acceptable to you:

Sign_____

- To translate English <EWB User—manual> in to Chinese version.
- Final report concerning above activities. The report will include:
  A discussion of the types of educational setting in which EWB is most like to have acceptance in our country, and why.
  What adaptations would have to be done or should be done for EWB to be acceptable in our country.
  The prediction of the educational value of EWB for our country, if the adaptations are implemented.

General prediction

EWB, this types of educational setting, will be well received in our country and can satisfy the Basic demands for the teaching of analog and digital circuits. Of course if components in the parts bin and instruments can be increased, the EWB should more convenient to satisfy the teaching demands.

We think present edition of EWB is feasible for the level of university, but to advance to use for vocational education and personnel training it is necessary to change several component' s signs and add a little Chinese characters into the EWB program which are more ease acceptable by all Chinese students and teachers.

We predict it is possible in our country to have over 1000 EWB users in coming two and three years.

591

THE ADDED VALUE OF NEW TECHNOLOGY IN COMPUTER ASSISTED SECOND LANGUAGE LEARNING

J.G.M. Jaspers, G. Kanselaar, & W.A.M. Kok, Department of Educational Research, University of Utrecht, Utrecht, The Netherlands

Abstract

The Intelligent Tutoring System *IT'S ENGLISH* is a computer–assisted learning environment for teaching foreign vocabulary by deducing meaning from the context, and by encouraging oral and written productions in the second language. To realize these ideas, *IT'S English* uses a syntactic parser and a database on CD–ROM, holding both auditory and visual information.

Research on the use of the program focused on three questions. First, how do learners use the different options offered by the program? Second, what is their appreciation of the program as users? And third, what are the learning results of the program?

Summary

Introduction to the problem

This paper presents a survey of the results of the project *The Use of CD–rom in Computer Assisted Second Language Learning* (SVO, 9104). In order to explore the potential of CD–rom technology in relation to second language learning, we developed a program for learning English vocabulary, called *IT'S English* (Intelligent Tutoring System for Learning English).

From September 1990 to January 1991 research was done with IT'S ENGLISH at two secondary schools. Research questions focused on the use of CD–ROM specific options in a fully learner controlled environment. Our research addressed the following questions:

- How do learners use the different options offered by the program? Is there a relation between the used options and certain types of exercises?
- What is their appreciation of the program as users?
- What are the learning results of the program?

Method

To answer the research questions we used the program *IT'S English* as a testing environment. *IT'S English is* an implementation of our theoretical starting points: the program fits in with a modern, communicative approach to second language learning, and takes into account teachers' criticism of existing courseware. More

specifically, the program is capable of handling productive language use, and encourages users to learn foreign vocabulary by deducing meaning from contexts. Furthermore, the program is flexible and teacher friendly: it allows teachers to add their own material without having to specify translations, context sentences, feedback, etc..

In order to realise these starting points, the program makes use of:

- a syntactic *parser*, capable of analyzing most English sentence structures used in the first years of secondary education. The parser makes (feedback on) free textinput possible.
- a monolingual *lexical database* on CD–ROM containing *descriptions* and *context sentences* in everyday English, and the grammatical features, synonyms and antonyms of over 70.000 words.
- a *database* on CD–ROM with the *digitized pronunciation* of 5000 words, 1200 context–sentences and a number of texts.
- hese databases can be consulted by the learner at any moment and thus constitute an extensive *help facility*.
- a *database of texts*. These texts can be taken from different teaching methods. They are the basis for several exercises.
- a *generator of exercises*. The program is able to generate exercises for both receptive and productive language use.
- *student registration*. Student performance is registered. This information is used for student specific exercises.
- t*ree generator*. Displays the constituent structure of input–sentences in the form of a tree.
- *feedback generator*. Generates feedback on student actions. The teacher can anticipate difficulties and add word specific feedback in a very simple way.
- *context sensitive help* on how the program can be used.

Data on the help options used by the students were measured in straight counts obtained from *IT'S English* logfiles. The students' appreciation of the program was measured by means of questionnaires. The learning results were tested in a pretest - post test - control group design. The independent variables were teaching methods (traditional versus communicative), type of school (havo, vwo) and experimental versus control group. The dependent variables were vocabulary test, cloze test, inference test and sentence production.

Results

Our first research question concerned the frequency with which pupils used different types of information about English words available on the CD–ROM. These different help–options have been used selectively by the pupils. In reading texts with new words they more often use the meaning definition of the new word, whereas in completing cloze texts they more often try to find the correct word by calling up context sentences in which the same word occurs in an everyday English context.

The assessment–scores for all kinds of user–aspects of the program, as well as its educational effectiveness, were mildly positive. The only aspect that showed a mildly negative score involved the speed of the program.

The results with respect to the last question (about the learning results) show that in the condition of the communicative teaching method the experimental group performed better on the vocabulary test.

Conclusion
As far as learning results are concerned, one of the conclusions is that learning words from contexts is a suitable method for more productive exercises. It still has to be shown that the acquired words can subsequently be used with greater facility in free writing.
From a 'new technology' point of view, one of the conclusions is that working with a network of XT–computers produces insufficient speed for a number of actions. Reaction–times exceeding two seconds quickly tend to become irritating.

Discussion
The practical value of CD–rom applications in the classroom will be discussed. Some general questions will be discussed: how do learners function in a totally self–controlled learning environment, is intelligent computer assisted language instruction feasible in classroom situations. Furthermore, attention will be given to the state of affairs with respect to the further development of *IT'S English*.

HOW FREE SHOULD STUDENTS BE? A CASE FROM CALL: COMPUTER ASSISTED LANGUAGE LEARNING

A.G. Sciarone, & P.J. Meijer, Applied Linguistics Unit, Delft University of Technology, Delft, The Netherlands

Computer programmes contribute in several ways to the teaching of foreign languages:

- By making language teaching less expensive. Programmes, in some cases, can be used to correct students' mistakes thus saving teaching time. Explanations concerning these mistakes can be fed back to the student along with further instruction. Grammar, vocabularies and explanations about exercises can be taken over from the teacher by the computer. By reducing the time required by students for correction of their mistakes and explanations of problems, the computer programmes can both save money and provide individual instruction. The extent to which computers actually make language tuition cheaper, however, is not clear: computers are very quickly out of date and development of software is expensive.
- Increasing Ease of Learning. Students can, to a large extent, decide for themselves, when and where they want to learn. Individual learning - in contradistinction to group instruction - is not hindered by the slower pace of other students. Explanations are given only when - and as often - as the student requires. The student becomes increasingly independent of his teacher and fellow pupils and so is more free in his learning. The extent to which computers can increase the ease of learning, as before, is not completely clear. Many programmes, especially those with multiple features and great flexibility require considerable instruction time. The instruction required for relatively simple word processing programmes is a case in point.
- Increasing the Learning Effect. The extent to which computer programmes can improve foreign language tuition in general, and the CALL system in particular, has yet to be clarified. This question forms the central theme of the discussion which follows.

During the past three years several experiments have been carried out to determine whether, and in which way, CALL improves learning a foreign language. Two questions were found to be important::

- How do students work when they learn a foreign language?
- Which learning tasks lead to the better results?

The first question is important as it determines the degree to which the tasks are being done. The answer is necessary as it provides the answer to the second question and so to the effect of CALL. The authors' papers on this subject have

been published in recent publications (Sciarone & Heerkens, 1989; Sciarone, 1990; Meijer, 1991; Sciarone & Meijer, 1991).
The students in these experiments, over a three-week period, were asked to learn three texts daily, of about 250 words. A text is judged to be well learned when the student listens to the text, or reads the text, and understands it completely without recourse to the vocabulary. With each text a gap-filling exercise was given having approximately 35 items. Unlike much practice material, the sentences were semantically related as in a normal text. Furthermore, various grammatical and vocabulary problems were practiced at the same time, the nature of these being selected on the basis of the text that had to be learned. To see the effect of such exercises (which are an important part of many language courses) on learning, the final test results of students who did and students who did not do gap-filling exercises during the course have been compared. The results in both cases are the same.

With gap-filling exercises			
n	k	x	s
81	300	192.7	57.7
(text)	85	60.2	17.7
(exer)	90	49.7	18
(new)	125	82.7	23.9

Without gap-filling exercises			
n	k	x	s
56	300	191.3	66.5
	85	56.1	19
	90	53.1	20.2
	125	82	27.6

(n = number of subjects; k = number of items; x = average; s = standard deviation)

In this context it should be mentioned that the exercise material used was not mechanical in nature and for that reason, could not be looked upon as ineffective (cf. Walz, 1984).
Before we can judge and put the blame on ineffectiveness of the exercise material, it is important to know how exercises are done. It is, for example, just possible that one might guess the answer and/or that parts are being skipped; in other words, that the material was not done well. To get a better insight we had students do exercises using the PC. All the items of the exercise were to be filled in; if an incorrect answer were given a message was displayed and the mistake had to be corrected. After three faulty attempts the correct answer was displayed by the programme and the student - after the provided answer had disappeared - still had to fill in the correct answer. At the end of the exercise the score was assessed and if a score were lower than 60% the exercise had to be repeated. For each student, a registration file recorded, amongst other data, the answers given and the clock time at the beginning and at the end.
This experiment showed:

- To do these exercises successfully, the students needed to put in more study time, far more than they were devoting to it at home;
- That many students had to do the exercises more than once because of inadequate scores;
- That many students, to save time, gave quick answers (which turned out to be 'empty' answers) in order to get the correct answer through the programme.The

exercise would have to be redone, but thanks to the stolen answers this could be done quickly and the exercise filled in practically faultlessly.
Conclusion: one can say that the freedom given these students during the exercises was especially employed just to reduce the exercise time, and that through this, the work was either not done at all or was done badly.
In the next experiment students again did exercises on the PC. This time, the students had to score at least 60% before the programme would provide the correct solution to an incorrect answer. But twice per exercise it was possible for the student to inspect which answers were good and which faulty. This experiment showed that many students were not capable of a score of 60%. Before seeking the cause in the exercise material, it was decided to find out how well the students had mastered the lessons that form the basis for the exercises.
Another experiment was therefore carried out in which access to the exercises was possible only when the students could show they had learned the text material satisfactorily since, naturally, that is the prerequisite for all other course activities. This material consists of the grammar, the vocabulary and the examples necessary for conversation and for generating the exercises based on all their contents. The experiment showed that students who learned their texts well would have no difficult with their exercises. This suggests that freedom with CALL is not effective for students who have not executed certain tasks properly.
On the basis of these findings, still another experiment was done which required mastering the texts from the course material. For this experiment a programme was developed in which the homework - the learning of texts - could be automatically examined. Each student received a different piece of text for which missing words had to be filled in. When an unsatisfactory score was made, the student was tested again on the same text, and so forth, until a satisfactory score was achieved. What was avoided was the usual practice of `punishing' the student with a low score then going on to the next text. To the contrary, they were forced to go back and learn the inadequately mastered text. Students who took part in this experiment achieved significantly better results on the final examination than the others.

	n	k	x	s
Experimental Group	65	100	77.2	14.7
Control Group	65	100	65.4	16.47

The final two experiments provided a partial answer to the second question: 'Which learning tasks lead to better result?' Increasing effectiveness was achieved by forcing the students to really learn the texts which, as has been shown, is the prerequisite for conversation and for doing exercises. Doing exercises, however, appeared to be not very effective. Students either showed through unsatisfactory execution of other tasks that they were not capable of doing them, or that they had no difficulty in doing them because they had done earlier tasks well. In either case the exercises were not really contributive.
As to the case of the behaviour of students during exercises, our findings tie in with the findings of Gardner, et al. Highly motivated students showed that they spent remarkably more time in learning a foreign language than less motivated students. Weak students tried everything to reduce their learning time as much as possible

597

through (mis)using the CALL facilities. Only after a reliable way was found of testing each student - and forcing them to spend more time learning when they got an unsatisfactory result - did the results improve. Our suggestion, therefore, is freedom of action for the CALL user to carry out learning tasks depending on the achievements s(he) has shown on learning tasks necessary for their successful completion.

References

Gardner, R.C., Smythe,P.C., Clément, R., & Gliksman, L. (1976). Second Language Learning: a social psychological perspective. The Canadian Modern Language Review, 32, 198-213.

Meijer, P.J. Zijn gerichte oefeningen het middel tegen de kwaal? (1990). Delfts Talen Bulletin, 1, 4-13.

Sciarone, A.G. Oefengedrag als factor in de effectiviteit van vreemde-talenonderwijs. In: B.R. Simons and J.G. Lodewijks. Onderwijsresearchdagen 1990. Technologie/Methodologie, 1990, 67-75.

Sciarone, A.G., & Heerkens, L.M.M. CALM. Computer-induced Adaptation of Learning Materials. In: E. Buchholz, Fremdsprachen mit Mikrocomputer. (1989). Wilhelm Pieck Universitat, Rostock.

Sciarone, A.G., & Meijer, P.J. Does practice make perfect? On the effects of exercises on second/foreign language acquisition. In preparation.

Walz, J. (1989). Context and contextualized language practice in foreign language teaching. Modern Language Journal, 73, 160-168.

DOES FEEDBACK ENHANCE COMPUTER ASSISTED LANGUAGE LEARNING?

Elisabeth van der Linden, Universiteit van Amsterdam, Amsterdam, The Netherlands

Programs for computer assisted language learning (CALL) are becoming more sophisticated: instead of confronting learners with simple drill and practice exercises, program designers are now developing courseware of a more open ended character (e.g. adventure type programs) or courseware providing extensive feedback to the learner. Developing this last type of courseware, in which feedback generally is based on an error analysis of a large group of learners, demands an enormous investment in time and energy. The question is whether this investment pays off: do learners indeed learn more by programs with a lot of feedback?

Research on the effect of feedback in L2-language teaching and learning in CALL is practically non-existent; even in "normal" L2 learning it is rare. In oral proficiency research, feedback is considered to be a part of "teacher style" (see Chaudron 1985), and nothing is clear about its effect on learning. Research on the effect of feedback on written L2-proficiency (Cohen & Robbins, 1976; Lalande 1982) shows generally that specific linguistic feedback does have some learning effect, but the results are rarely significant. Generally it is stated that the effect of feedback is difficult to measure. This is all the more so in written language production, where feedback never comes immediately after the production of text, a fact that diminishes its impact.

It is here that we find an important difference with computer aided instruction, since this is characterized by immediate feedback to students' responses. Many CALL-experts tell us that this is what makes CALL (and more generally CAI) so effective (e.g. Moonen & Gastkemper 1983, Higgins & Johns 1984, etc). In studying the effect of feedback we have to find out in the first place how students react to feedback that is offered to them: do they pay attention to it and does it help them to learn?

A project at the French Department of the University of Amsterdam is trying to find out how effective feedback really is or can be by studying closely the behavior of students working with a series of grammar learning programs. In this project, students are confronted with traditional grammar problem-solving tasks (see also Breuker & van der Linden, 1983; van der Linden, 1989): a grammar problem is presented to them that they have to solve. In giving their response, they can choose to make use of "hints" in the program. After giving their solution to the grammar problem, they are offered specific detailed feedback about their

response. If their response was not correct, they are allowed as many new tries as they want. Each try will offer them detailed feedback in the case of erroneous responses. They can equally look up the right answer whenever they want. This answer is not given atomatically after a maximum number of tries, like in other programs.

When attending to the CALL-sessions in the CALL lab, the strong impression arose in the research team that students profit in very different ways from the information given to them in the feedback. Some seemed even to overlook the feedback entirely, interested as they were only in finding out the right solution to the grammar problem. Two methods of studying students' strategies in the CALL lab were employed. First, log files were studied closely containing the series of tasks - responses - feedback of each individual student. In addition to that, thinking aloud protocols were analysed. The results of the log files show rather large individual differences in the preferred strategies of the students. Some students seem to read indeed the feedback very carefully. They go on trying to find the right answer, using up to eight or ten trials per sentence when necessary. Thisis the group that works more or less like we expected.Another group of students seems not to be interested in the feedback: they do not correct themselves when the computer tells them they made a mistake. Probably, they don't read the feedback: they "jump" immediately to the right answer. After looking this up, they go on to the next sentence. Some students go even further: they do not look up the right answer when they are told they made a mistake but just go on to the next sentence. It looks as fi they "scan" the program in order to find out what percentage of the sentences they did right the first time. The first results of the protocol analysis seem to confirm these findings. Student strategies present important individual differences. For this (rather small) group of learners, there is no clear relationship between the chosen strategy and the learning effect. The protocol analysis however has not been completed for the moment. Detailed results will be presented at the conference.

References

Breuker, J., & van der Linden, E. (1983). Het oplossen van Franse grammatikaproblemen. In: J. Beishuizen e.a.Onderwijsleerprocessen: 113-123. Lisse: Zwets & Zeitlinger

Chaudron, C.A. (1985). Comprehension, Comprehensibility and Learning. Studies in Second Language Acquisition, 7, 2, 216-232

Cohen, A., & Robbins, M. (1976). Toward assessing interlanguage performance: the relationship between selected errors, learners' characteristics and learners' explanations. Language Learning, 26, 45-66.

Higgins, J., & Johns, T. (1984). Computers in Language Learning. London: Collins.

Lalande, J.F. (1982). Reducing composition Errors: an experiment. The Modern Language Journal, 66, 2, 140-149

Linden, E. van der (1989). Feedback in het computerondersteund talenonderwijs. Toegepaste Taalwetenschap in Artikelen, 33, 64-71

Moonen, J., & Gastkemper, F. (1983) Computergestuurd onderwijs. Utrecht: Spectrum.

ANALYSIS CRITERIA FOR THE EVALUATION OF DECISION SUPPORT SYSTEMS (DSS) USED BY PUBLIC EDUCATIONAL AGENCIES (PEA) IN VOCATIONAL GUIDANCE (VG)

Giuseppe Mantovani, & Mirco Bolzoni, Department of General Psychology, University of Padova, Italy

We are interested in monitoring the impact of DSS on educational environments, thus our attention is directed towards applications of DSS to VG and to appropriate criteria of analysis for the new systems. A suitable approach is on focusing the attention on both the overall scenario where the interaction takes place (cfr. Carroll, 1991) and the design process (cfr. Mantovani, 1990, 1991a). We considered three DSS applied by several PEA for VG purposes in Northern Italy: type–1 (ARIANNA 1) has been developed and is presently used in Turin, type–2 (ORIENTESIS) in Bologna, type–3 (MITO) in Milan and other places. The methodologies we adopted were semi–structured interviews with social actors involved in the implementation and use of the systems, direct observation of users' behaviour, follow–up interviews with users.

Firstly, we realized that intelligent systems for VG are born and applied in quite different social and organizational contexts. They differ in aim, structure, design philosophy and social setting of application. The relevance of the contextual aspects of the introduction of VG/DSS in the VG process is observable in the break up of roles that are distributed within different actors. So we can see that the VG USER (who is the user of the VG process), the SYSTEM USER (who is the user of system's performance), and the SYSTEM INTERFACE (who is the human interface to the system) may not coincide with the same individual rendering the harmonization of tasks and goals problematic. The same roles considered in the VG/DSS design perspective prove to be relevant in the search for the appropriate model of the user: VG USER is the content of system's information; SYSTEM USER is the manipulator of the information outcome of the system; SYSTEM INTERFACE is the ergonomic interface to the system. Three models of the VG process were identified corresponding to the VG/DSS types. The Peripherical Adoption for type–1, the Mediated Adoption for type–2 and the Direct Adoption for type–3 (see Figures 4,5,6). These models expose the different emphasis posed on the participation of the job–seeker to the process and the number of necessary steps involved in the channeling of information (Distance of the interaction, as shown in figure 7). Afterwards, we built a social cues typology (SCT, see figure 1),

taking into account how roles and actors are connected for each system type. This table offers clues for the "weighting" of the different models of VG User, System Interface and System User, handled by the VG/DSS.

Secondly, we performed a structural analysis of the features of the DSS as suitable dialogue tools. We emphasized the role of the quality of the human–computer interaction, that seems to be thecritical factor in the success of the dialogue. That would require the construction of an appropriate model of the user, but as contacts are placed by various actors, within different roles, the design process remains problematic. The design task is demanding because each individual brings specific expectations and communicative skills and the system has to attune its communicative skills with these various actors. This analysis was formalized in a typology of the system interaction capacity (SIC) as shown in figure 2. The critical variables of the SIC table are: symmetry of dialogue, reciprocity of response, cooperation on tasks, synchrony of response.

Thirdly, we analyzed the users' responses to the different DSS in the different PEA's social contexts. The user demand and motivation towards his VG need is often ambiguous. This aspect stands out particularly in the design process where designers and PEA officers come to determine the needs of their hypothetical user with a strident accuracy compared to that brought by the real user. The result might consist in a mismatch of goals and tasks of the VG/DSS driving the consulting session to the fulfillment of non existing needs. Two resolving strategies are available to the design process: one is in increasing the distance of the interaction with the improvement of the participation of human actors in the VG process; the other is in enforcing the cognitive potential and communication standards of the VG/DSS in order to cope with the hazy needs of the VG user. Second strategy is recommended when the distance of the interaction is low and becomes a must when the human support to the interaction is absent. The interactive quality of the DSS will enable the user to control the development of the counseling session according to his/her needs. Therefore, we built a typology of the dialogue style (DST) fostered by the three DSS (see figure 3). Critical variables for the DST are: cognitive enforcement messages, interpersonal addressing, graphic/animated interface, tutoring style, natural language, use of commands. Connexions between SCT, SiC and DST allow a widened perspective for a proper DSS's design and implementing strategy according to each VG process setting.

References

Carroll, J.M. (ed.) (1991). Designing Interaction. Psychology of the Human Computer Interface. Cambridge University Press.

Mantovani, G. (1990). Usabilità: compatibilità cognitiva ed equilibrio nel controllo tra progettisti ed utilizzatori, in "Sistemi Intelligenti", 4, 243–268.

Mantovani, G. (1991 a). La qualità dell'interazione uomo–computer. Controllo, responsabilità, modelli nei sistemi interattivi. Il Mulino,Bologna.

Mantovani, G. (1991 b). Expertise and Intelligent Training Systems in Work Environments. Paper for the AIOSP/IAEVG (International Association for Educational and Vocational Guidance) International Conference, University of Lisbon, September 9–13.

figures

Actors and roles for DSS type	VG USER	SYSTEM USER	SYSTEM INTERFACE
Families	Type-1	None	None
School teachers	Type-1	None	None
People looking for advice, kids	Type-2 Type-3	Type-3 (Type-2)	Type-3
Job-seekers	Type-2 Type-3	Type-3 (Type-2)	Type-3
PEA-Officers/VG Tutors	Type-2	Type-1 Type-2	Type-2

Figure 1. Social Cues Typology. Connections between actors and roles of the VG process for DSS type.

SIC	Type-1	Type-2	Type-3
Symmetry	yes	no	yes
Reciprocity	no	no	yes
Cooperation	yes	yes	yes
Synchrony	no	no	no

Figure 2. System Interaction Capacity for DSS type.

DST	Type-1	Type-2	Type-3
Cognitive enf. mess.	no	no	yes
Interpersonal Addr.	no	no	yes
Graphic/Animated Int.	yes	no	yes
Tutoring Style	yes	no	yes
Natural Language	yes	no	yes
Commands use	yes	yes	yes

Figure 3. Dialogue Style Typology (DST) for DSS type.

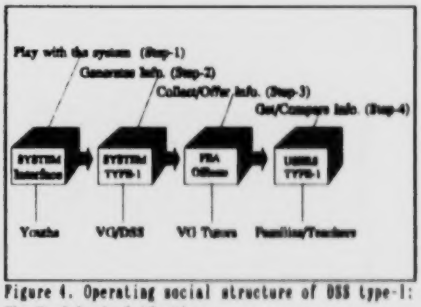


Figure 4. Operating social structure of DSS type-1: The Peripherical Adoption.

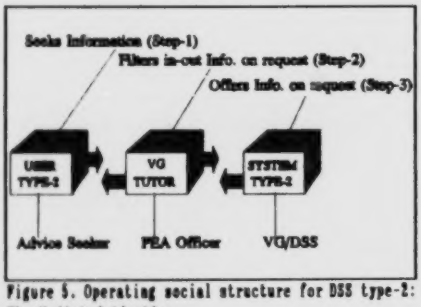


Figure 5. Operating social structure for DSS type-2: The Mediated Adoption.

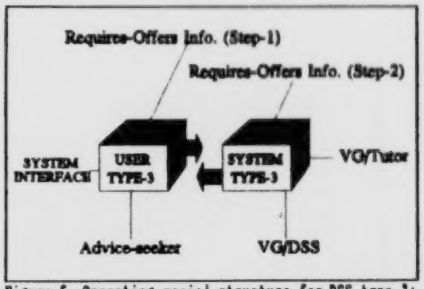


Figure 6. Operating social structure for DSS type-3: The Direct Adoption.

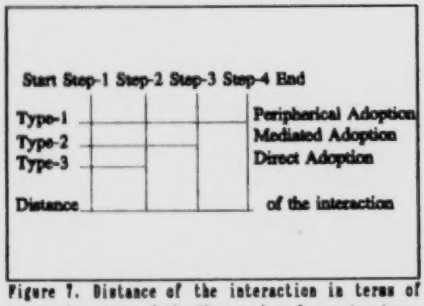


Figure 7. Distance of the interaction in terms of processing steps of the VG procedure for system type.

BEST COPY AVAILABLE 603

MULTIMEDIA IN INDUSTRY AND EDUCATION: A DECISION MODEL FOR DESIGN

Lynne Reynolds, & Diane Ehrlich, Northeastern Illinois University, USA

The need to develop life-long learners has become increasingly urgent as we move into the twenty-first century. Technology and the move toward a global society demand continuous adaptation and change. Schools are under pressure from society to create a quality curriculum with additional components which go far beyond that which was expected fifteen or twenty years ago. In order to remain competitive industry is required to train individuals to work in particular environments and to upgrade skills and knowledge as new developments unfold. Fortune Magazine states that the most successful corporation of the 1990's "will be something called a learning organization, a consummately adaptive enterprise."

Traditional methods of delivering training and education have had limited success in meeting the fast-paced and high-quality demands of learning environments today. Attracting and retaining culturally and academically diverse learners in schools today will be strongly influenced by the effectiveness of instructional delivery systems and whether or not learners can both relate to the material presented and see an almost immediate application for their learning. Multimedia can change the look and feel of learning by providing an opportunity to reach people with different learning styles, different skill levels, and in different geographical areas; multimedia offers the potential to reduce the learning curve and accelerate the learning process. Rather than being seen as expensive and frivolous, multimedia approaches may then be seen as an economic necessity both in industry and education. To be sure, there are situations where traditional approaches may actually be the preferred delivery system. It is also possible to become so enamored of a particular technology that the equipment itself becomes more important than the need for learning. In order to make an informed decision about the use of multimedia, a systematic approach to design should be firmly in place and it should have as its driving force the needs and goals of a particular learning situation.

Figure 1 is a graphic representation of the design process in a puzzle form comprised of a series of interlocking pieces. The depiction of the process is not linear because of the reiterative nature of the process which may begin at almost any point. However, to be effective the process must be driven by clearly identified needs and goals, as with any good design model. This does not mean that the

needs and goals may not shift as different pieces of the process are introduced, such as rapidly developing technology or budgetary constraints. The design model described here is derived from traditional instructional system design models (Kemp, Dick & Carey, Romiszowski); however, the potential impact of multimedia as a delivery system/learning tool was not often addressed as part of this process.

It is apparent when looking at the possibilities multimedia has to offer that communication with experts in other disciplines will be vital to the success of a project. The very nature of multimedia creates a need for interdisciplinary dialogue and a blurring of discipline lines. It is not only educators, trainers, performance technologists, and instructional designers who are interested, but computer programmers, video and audio production specialists, and marketing professionals who are actively involved. Members of this diverse team must have a common understanding of the instructional design process to make the best use of each unique orientation to multimedia. As each element in the model is addressed it is important to note that previously developed elements may need to be evaluated and then revised in light of new input.

In the analysis phase of design a consideration of the problems that need to be solved through instruction should be systematically developed. The needs and goals may change as a project takes shape because the primary concern in this phase is client-centered. In both school and organizational settings, the client's expectations are key to the success of the end product.

Although correct analysis is basic to effective instructional design, many clients by-pass this stage and assume that they know the solution. With the high costs of developing technological instruction, it is imperative to conduct an analysis of the organization, the task, and the learners before any action is taken.

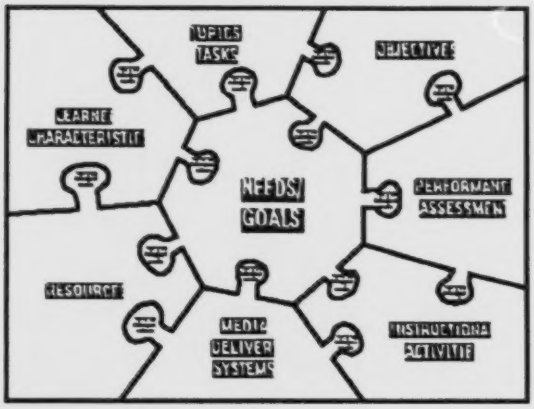


Figure 1
Decision Model for Multimedia Design

605

Learner characteristics reflect the demographic attributes, learning styles, readiness, and motivation to learn of the target audience. Variables such as learners' expectations affect the amount of time is needed for instruction, the level of instruction, and the varied approaches that are required.

Topics and tasks are identified by performing a content or task analysis to determine what components should be included in instruction. Subject matter specialists or exemplary performers can provide the specific areas to be incorporated. At this stage of the design process, heavy emphasis is placed on researching items to include in the content and finding a sufficient number of sources to support the level of detail to provide the foundation for instruction.

Objectives are directly related to the goals and reflect the intent of instruction, and should focus on learning outcomes and should be stated in measurable terms. While all members of the team should be consulted and involved in the development of project objectives the refinement of those objectives may best be left to the team members who have this expertise.

The methods of ***performance assessment*** should match the desired outcomes of instruction. They may range from simple pencil and paper tests to working through complex case studies. Media may be incorporated in assessing performance so that the learner may measure his/her own performance and/or review it at a later date. According to Kirkpatrick (1976) performance assessment should occur on several levels: 1) reaction; 2) learning; 3) changes in behavior; and 4) long-term results, taking place immediately following instruction and also several months after instruction. Multimedia can play an important role in post-instructional assessment through teleconferencing, distance learning activities, etc.

Instructional activities are the learning experiences developed to present instruction and to allow the learners to demonstrate their ability to meet the desired performance level. Activities may range from delivering instruction via an instructor dominated mode (such as lecture) to delivering instruction through a vehicle such as multimedia where many methods may be woven together. Activities need to be carefully planned and controlled by the allotted time which will determine the depth of the content and the number of activities which can be designed to support learning.

The choices for ***media/delivery systems*** for instruction are extensive. If one is limited to a traditional classroom model, then lecture may be the primary mode of delivery. However, with the range of media available to entice, enhance, and reinforce learning, the selection of media/delivery systems may be one of the most important elements in designing instruction. The integration of instructional design and multimedia will be a key factor in the learning process of students in educational systems today because of the many changes in the instructional process.

Resources impact all stages of the design cycle. Limited resources may lead to the conclusion that the use of multimedia is not a viable option. However, the use of technology is cost-effective when it is selected based on strategic considerations. Many instructional designers are hesitant to work with technology because of a lack of knowledge and experience. For this reason many companies are using teams comprised of project managers and administrators, graphic artists, instructional designers, instructors/teachers, media specialists, writers, subject matter experts, and performance technologists whose combined expertise

produces a more integrated learning experience. Many specialists are used to working in isolation or have only been involved at certain points during the developmental process with other types of projects. Instructional multimedia requires a different approach. Team members may be unaware of the tasks for which other members are responsible and lack appreciation for the amount of time these tasks realistically take. This necessitates that time be spent on building a psychological support structure to ensure the success of the team. Project management skills are crucial so that one member of a team is not wasting time while another fails to meet a deadline. All team members should be aware of budgetary limitations, completion dates, and the most effective distribution of resources. A Project Task Timeline has been developed and lists a number of tasks that should be completed as part of the design process, the most important being the development of clearly defined instructional goals which leads to clearly defined roles and responsibilities for team members.

The entire process goes through continuous ***revision and evaluation***. Each decision made influences many others so that the various elements are constantly being reviewed. A Checklist for Multimedia Design has also been developed which suggests questions that the instructional design team should ask. The form follows the pieces of the multimedia decision model and provides a series of items which to discuss so that all members of the team become aware of the broad scope of the project and the variety of decisions which are made during the stages of the design process. The questions are designed to encourage a sharing of ideas and reservations and also to provide team members with a common vocabulary. Specific tasks should be assigned to team members as the questions on the checklist are discussed, or at a later date. When the team invests the needed time up-front to systemically research the pertinent issues, team members acquire a wider perspective.

The increased number of participants attending international conferences demonstrates that as we move to a global economy, nations of the world share many common problems. An assessment of the educational system in the United States is documented in two reports: America 2000 and the Secretary's Commission on Necessary Skills (SCANS) report. The latter report ties together the importance of instructional design, multimedia and the need to change the way education is structured. Central to the development of an agenda for the use of technology is a systematic assessment of needs and goals. This requires the development of teams to ensure a comprehensive understanding of what technology and multimedia can bring to the educational process rather than being seduced by its impressive capabilities. The impact of technology on instruction cannot be denied. However, the message and not the media/delivery system is the central issue. The design model encourages a total integration of both the message and the media/delivery system.

References

America 2000: An Education Srategy. (1991). Washington D.C.: U.S. Department of Education.

Bigge, J. (1988). Curriculum based instruction for special education sudents. Mountain View, California: Mayfield Publishing Company.

Bloom, B.S. (1968). Learning for mastery. Evaluation comment, Vol 1, No 2. California: UCLA.
Dick, W., & Carey, L. (1990). The systematic design of instruction. 3rd Ed. Glenview, Illinois: Scott Foresman & Company.
Dumaine, B. (1989). What the leaders of tomorrow see. Fortune Magazine. July 3, 48-62.
Kemp, J.E. (1985). The instructional design process. New York: Harper & Row.
Kirkpatrick, D. (1975). Evaluating training programs. Madison, Wisconsin: American Society for Training and Development.
Landa, L.N. (1974). Algorithimization in learning and instruction. Englewood Cliffs, New Jersey: Educational Technology Publications.
Romiszowski, A.J. (1974). Selection and use of instructional media: A systems approach. London: Kogan Page.
Romiszowski, A.J. (1984). Producing instructional systems: Lesson planning for individualized and group learning activities. New York: Nichols Publishing.
Secretary's Commission on Achieving Necessary Skills (1991). What Work Requires of School: a SCANS Report for America 2000. Washington D.C.: U.S. Department of Labor.

TRAINING LIBRARY-USERS IN SEARCHING LITERATURE SYSTEMATICALLY WITH COMPUTER ASSISTED LEARNING: THE FIDELITY OF A SIMULATION

J.Braaksma, & A.Pilot, University of Twente, Enschede, The Netherlands

The research reported in this paper concentrated on the use of a certain form of simulation in a CAL program, and more specific on the parameter 'fidelity' (Alessi and Trollip, 1991). Fidelity refers to how closely a simulation imitates reality. There is a complex relation between transfer of learning and the fidelity in an instructional setting. The level of fidelity should be appropriate during the whole learning process of the student.

The parameter 'fidelity' is often an important one when designing CAL. This research tries to enlarge the knowledge on this issue and is part of the evaluation of the use of CAL in learning to search literature systematically.

The aim of the course in which the CAL-program is used is training library users in searching literature systematically. Therefore the student is trained to use a procedure starting with a rather open formulated problem statement. The scope of the simulated task is to study the literature regarding the problem and to produce a report. The skills that should be learned are: (re)formulating and clarifying fuzzy presented problem statements, choosing terms for searching in sources, the actual searching in different types of sourses, applying a useful way of documenting results, choosing the optimal sequence of steps in their work etc.

Many parts of training searching literature systematically can be realised by written material, through verbal presentations etc., but some parts (like changing attitudes and the training of skills) should be trained by exercises and intensive feedback. Traditionally this kind of instruction is judged as very time consuming by the teachers and as boring by the students.

Using Computer Assisted Learning as one of the instructional means in this course might well change these negative judgements in more positive ones. This is especially expected because of the possibilities of CAL to give frequently feedback of high quality during the learning process.

In the CAL program used in this research-project two alternative forms of lay out can be chosen. The first alternative is to use pictures of real documents by scanning these documents. The second alternative is to type these documents again with the editor of the CAL authoring system, and use this mode of presentation in the courseware.

Generally the advantages and disadvantages of these alternatives seem to be that the readability of the typed solution is higher. The fidelity is higher in the scanned

solution including all graphical details, lay out, meaning and alternation of lettering and other signs in bibliographical sources. The effort to produce scanned materials is higher, as are the costs. The disk space needed and technical problems such as graphic cards are also unfavourable when scanned sources are used.
Because of the expected learning effects and satisfaction of students, it was decided to scan most of the sources. However, for future versions of the program in other languages or with other cases it is important to know whether or not the assumed surplus value in terms of satisfaction and learning results can be established.

Design of the experiment
For this experiment 60 students used the CAL program in an experimental and controlled situation. They worked in groups of two. Fourteen groups got the scanned source first; 16 groups got a typed source first.
Their instructional process is evaluated in detail with logfiles, observations, written tests and questionnaires: how do they examine the program precisely, especially the screens with the texts mentioned. The students were also interviewed about the fidelity of this part of the course. Their learning results were evaluated using a written test after their work with the program. In that test they had to solve a similar kind of problem as was trained in the CAL program.
To facilitate a comparison between typed and scanned texts, two secondary bibliographical sources were produced either typed or scanned. After the first and after the second source and at the end of the CAL program the students answered a questionnaire.

Results
The students worked as planned. They answered all questions in the program and did all tasks according to the design. They needed about 2 hours for the CAL-program; the whole course took about 25 hours of study time.
In the paper quantitative data will be presented; in this summary attention is paid to qualitative data.

Time spent on learning
The logfiles show no significant difference between the time needed for watching the typed and scanned texts of the first bibliographical source. Watching the first source too about three times the time needed for watching the second source. Therefore it is concluded that the time spent consulting the bibliographical sources presented in the program is mainly regulated by the content of the sources and not by their representation.

Intensity of study activities
The observations of the students while going through the CAL program show that the first source is watched more intensively when it is scanned than when it is typed. The groups which got first scanned texts take more notes regarding titles and paragraphs than the others. This applies for the first (scanned) and for the second (typed) source presented.

The groups which got typed texts first restrict their note taking more to titles only. These groups also pickup more detailed information about the titles presented than the others.

Judgement by students

The data collected through the first questionnaire show that the group which first got a typed version judged them as more clear than the others. They also thought the example presented being a good one for searching a real printed bibliography. On the whole the opinions of these groups were less diverse than the opinions of the group which first got the scanned texts.

The second questionnaire was taken after the (now more experienced!) students also studied the second source. The group which had had the scanned source first and then the typed one judged the typed one as clearer. The group having had typed texts first judged them as clearer than the scanned ones but this group also experienced the scanned texts as being more representative for reality than the typed ones.
On the whole all groups were more explicit about the advantages of typing texts and about the disadvantages of scanning texts than for the advantages of scanning and the disadvantages of typing.
At the end of the CAL-program the students evaluated the CAL program as a whole. It is regarded as a very useful and rather pleasant one. They found it not too hard to go through but needed the teacher's assistance every now and then.

Learning results

While going through the CAL program students had to make notes. These notes were checked on some aspects that are especially relevant for the fidelity issue, like quality and completeness. It is remarkable that the groups which first got scanned texts made better notes than the others.
Students had to make a written test later during the course as indication for the effectiveness of the CAL program. All students turned out to perform sufficiently on this test.

Conclusions

Scanned texts turned out to have a high fidelity and to be valuable for learning effects in this training program for searching literature systematically. They are judged by students as valuable representations of reality. However scanned texts are also judged as less clear than typed texts. Therefore the overall conclusion is that typed texts (low fidelity) can and should be used when scanned texts are not necessary. This implies that the first training might well be most effective with typed texts and that illustrations and exercises should be presented as scanned texts. This fits in with Allessi and Trollip's theory on fidelity.

References

Alessi, S.M., & Trollip, S.R. (1991). Computer-based instruction: methods and development. Englewood Cliffs: Prentice Hall.

HYPERMEDIA FOR EDUCATION: THE CONTROL OF INFORMATION

Sjoerd de Vries, & Piet Kommers, University of Twente, Enschede, The Netherlands

Introduction

The use of Hypermedia Oriented Courseware Products (HOCP's) in education offers a wide range of opportunities for the teacher and the student. Students can for example use HOCP's to explore, and construct knowledge from information resources and HOCP's can support the learning processes. But the use of HOCP's in education is hindered by problems related to the control of information. We consider the environment for the use of HOCP's (ECU-I = Environment for Courseware Use - Inquiry products) as a framework for understanding and unravelling the problems related to the control of information. An important characteristic of an HOCP is the freedom in completing tasks. This freedom for example in exploring information resources leads to the need of a dedicated set of computerized tools to keep the teacher or learner as users of an HOCP in control of the retrieved information . If we don't realise an applicable set of tools, the use of HOCP in education will be poor and even detrimental to the role of the teacher.

Method

HOCP's can be categorized to inquiry oriented instruction '... a view of students as active constructors of knowledge, as knowledge as open and evolving, of academic learning as exiting and vital, and of teaching as a stimulus to curiosity and a model of inquiry' (Cohen, 1988). A courseware product enables Computer Based Education (CBE). A courseware product consists out of (1) a computer program that presents educational material to a learner and facilitates individual interaction between the learner and the computer and (2) if necessary additional materials to support the instructional process. A HOCP is a courseware product that makes use of functions provided by a Hypermedia program.
The concept of Hypermedia refers to the concepts of Hypertext and Multimedia. A Hypertext is a non-sequential writing, it is a directed graph, and each node contains some amount of information. The nodes are connected by directed links. Users read information, for example text, in the nodes and they navigate between the nodes by following the links. Multimedia refer to a collection of at least two different, but related information sources, like text, graphics and audio. At least one of the sources must be presented by means of a computer and the user must be able to act interactively with the information sources. Hypermedia in a technical

sense can be described as a computer-based approach to information management in which data are stored in networks of nodes connected by links. The nodes can contain text, graphics, audio, video, e.g. From the perspective of a user, Hypermedia programs are knowledge bases in which every display provide access to others. A knowledge base, as used here, refers to a data base containing primary information instead of for example bibliographic references. The goal of the use of Hypermedia programs in education is to make large multimedia storage systems accessible and utilise the systems to match curriculum and classroom needs. In this way the use of Hypermedia programs is directed to the retrieval and processing of knowledge by the user. A large multimedia storage system in this context can be seen as a heterogeneous knowledge base. Conventional methods of information management though are not sufficient for organizing and accessing a rapidly expanding mass of data.
At least three functions of Hypermedia programs can be distinguished. To begin with, Hypermedia programs offer the user options to inquire and explore knowledge bases in an interactive way. Secondly, Hypermedia programs enable the user to organize and construct knowledge, this means adding, deleting and editing options. Thirdly, Hypermedia programs support the processing of the information from the knowledge bases by the user by built-in options and features, for example by means of instructional questions, or special delivery methods of information.
Three parts of a HOCP can be distinguished. The first part consists of the learner goals, which describe the desired learner effects. The second part consist of the instructions that will be presented to the learners. Instructions for example may consist of directives and guide-lines. The learner goals and the instructions can consist of printed material. The third part is the knowledge base. The Hypermedia program is used as a tool to execute tasks related to the knowledge base and to the instructions. It is clear from the description of the functions that a HOCP is an adaptable courseware product. An adaptable courseware product is a product that can be adapted by users in accordance with specific conditions of use. The knowledge bases for example can be edited by the users and also the instructional content for the student can easily be altered by teachers.

We mention three expectations of the use of Hypermedia programs and related multimedia storage systems in education.

- Hypermedia programs offer teachers and students an open and flexible access to large multimedia storage systems. This means that the user is in control of performing tasks by means of the computer program, the user has the freedom to decide what to do. The user has to develop methods of searching and the computer can support the user in the process of information retrieval;
- A Hypermedia program offers the opportunity to tailor the HOCP to the users. All the three parts of a HOCP can be tailored. Hypermedia programs for example offer the possibility to adapt the information in large multimedia storage systems. At least three arguments can be mentioned for the use of adaptable courseware products in education. (1) The user base of a product can be widened. (2) The life-span of a product can be extended. (3) Product characteristics can be adapted to improve educational effectiveness;

- Finally, the function of supporting the information processing by the user from the multimedia storage system can be considered as a way of increasing the effectiveness of the use of those systems in education.

Four problems of the use of Hypermedia programs in education are the following.

- The usability of Hypermedia programs is not always considered as satisfactory. The usability is the extent to which an end-user is able to carry out required tasks successfully, and without difficulty, using the computer application system. A problem related to the usability for example is the orientation and navigation in the multimedia storage system;
- The students as users may lack the cognitive skills and dispositions required to take advantage of the functions offered by Hypermedia programs, for example skills and dispositions like: (1) some learners are in the habit of accepting information passively or of consistently seeking external assistance, (2) many individuals have trouble adjusting to self-directed learning, they lack effective study strategies and skills as well as basic responsibility skills related to self-management, and the motivation to learn;
- The effective use of large multimedia storage systems demands new information-seeking strategies. A clear understanding of those strategies needs further research;
- The function of adaptability of the multimedia storage system by the users leads to the problem of maintenance, and sustaining the functionality of the system. The problem of the maintenance of the multimedia storage system arises for example if the user wants to add information. The functionality of a multimedia storage system is hard to define. An important feature of the functionality is that the links that exist in the system must represent a clear relation, for example 'part of', 'example of', 'is caused by'. Users adapting the system can interpret links in different ways, so that the next user, who tries to retrieve information from the system has difficulties in finding correct information.

The four problems mentioned above can be related to the concept of information control. The problem of usability indicates that the use of HOCP's is constrained by a lack of control. The problem of the students who are not capable of making optimal use of such a product can be due to the usability but also to the functionality of the Hypermedia programs. Effective use of these products is a function of built-in information-seeking strategies. The problem of maintenance, and sustaining the functionality of the system is a problem of the control of information.

The environment for the use of HOCP's (ECU-I) is presented as an approach for modelling the information. De Vries & van Schaik (1992) described an ECU as an information system, because (1) the information and the flow of information is of great importance in an ECU, and (2) an ECU has to offer the users the possibility to be in control of the information. An environment for courseware use (ECU) is defined as a collection of tools to carry out activities related to the use of adaptable courseware products. Based on the functions of Hypermedia programs, HOCP's can be considered as adaptable courseware products. Based on the information model we have designed a prototype of an ECU for HOCP's (ECU-I). We have

used the principles of TEKSTNET and HYPERNET (Kommers, 1991) to realise the prototype.
We have carried out a formative evaluation of the ECU-I using the illuminative model of evaluation. The ECU-I was used in a research project for individual learning in two experimental schools. The ECU-I is used in the school years 1989-1990 and 1990-1991 during courses of three to six lessons in the first and second classes. At level 1 students used the ECU-I to explore conceptual schemes, and use the schemes while learning from texts. At level 2 students have used the ECU-I to search for information necessary for authoring an essay. We participated in the lessons and observed the teacher and the learners working with the ECU-I. Afterwards the teachers were interviewed about their experiences with adapting the HOCP's and using them in the classroom. The students responded a questionnaire about the use of the ECU-I. The data were analysed in a qualitative and quantitative way.

Results
Results will be presented about the use of the ECU-I during two phases in the use cycle of HOCP's: Adapt before use, and Use. In the phase Adapt before use the teachers adapt the HOCP according specific conditions of use. The three distinguished parts of the HOCP can be adapted. In the phase Use the HOCP is used in the classroom. The following results will be presented. (1) The perception of the teachers about the usability of the ECU-I in relation with the concept control of information. (2) The opinions of the teachers and the students about the usability, and the perception of the teachers and students about the learning effects of the HOCP's.

Conclusion
The freedom of the users in completing tasks related to the use of HOCP's can be balanced by the use of an ECU-I. The ECU-I can help to keep the users of HOCP's in control of information.

References
Cohen, D.K. (1988). Educational Technology and School Organization. In Nickerson, R.S. &Zodhiates, P.P. (Eds.), Technology in Education: Looking toward 2020. Hillsdale: Lawrence Erlbaum Associates.

Knussen, Tanner, & Kibby (1991). An approach to the evaluation of Hypermedia. Computers & Education, 17(1), 13-24.

Kommers, P.A.M. (1990). Hypertext and the acquisition of knowledge. Dissertation. Enschede: University of Twente.

Vries, S.A. de, & Schaik, P. van (1992). Aspects of an environment for the use of adaptable courseware products. In De Diana, I.P.F. & Schaik, P. van (Eds.) Aspects of courseware engineering. Berlin: Springer.

GENERIC MULTIMEDIA DATABASE: NEW DIDACTICS

R. Bestebreurtje, & P.W. Verhagen, Department of Education, University of Twente, Enschede, The Netherlands

Introduction: background

Economical factors lead to (some) use of optical media ("multimedia") in educational settings as generic multimedia databases. Use by multiple target groups then leads to a reduction of costs. However, from a didactical perspective these generic databases have limited applicability. Although use as a (hypertext) encyclopedia may be of value, for instance for assignments for which information retrieval tasks are essential, Dutch secondary and vocational schools do not often need this type of use. Instead the educational tradition emphasizes systematic traininglike instructional methods. Generic multimedia databases of themselves lack the instructional structures that are needed for this.

The current research project "Instructional Design of an Optical DataBase (ODB)" started in 1989 at the University of Twente. In this project two approaches -- construction of generic databases and the instructional design of interactive video- - come together, in order to develop a mutually enriched combination.

The project

The ODB-project is an attempt to develop the didactics (instructional principles) for the use of multimedia databases for multiple educational targets.

Central position with respect to the access of the learners is the teacher: The teacher is seen as the manager of the instruction who defines, organises, and evaluates the access of each learner to the database.

The types of application that are supported by the system are:

- Instruction. The type of instruction is related to the type of knowledge and the required level of performance. This is the major access to the database. It is organised by the construction of default "standard instruction route" for each target group. Each standard instruction route is constructed according to the greatest common divisor of the curricula of the schools of each level;
- Use of the database as a hypermedia encyclopedia, partly additional to the instructional information;
- Classroom-presentations. Use can be made of all types of information that are available for instructional and encyclopdial use, to construct a dedicated classroom-presentation;
- Access to the learner registration files (for the teacher only).

The product: The "Optical DataBase"
The multimedia product of the ODB-project will be the "Optical DataBase", for use in Dutch vocational schools. A typical Dutch domain of knowledge is chosen: "cheesemaking". The size of this knowledge domain is limited, which makes it practical for the experiments. It is taught at all levels of the agricultural education: lower-, middle-, higher agricultural schools, and at the agricultural university. Besides, it is part of the extensive industrial training system at all levels. The target groups of the ODB-project lie in principle at all these levels. As a first approach the project aims at the lower and the two middle agricultural schooltypes. Cooperation exists with six of the total of eight of these schools that have "food technology" as a body of study.

The target of the product
The project is directed to provide the teacher with tools to adjust the standard instruction route in accordance with local needs. It is possible for the teacher to:

- switch on/off specific topics;
- change the educational level of treatment;
- change the form of appearance by switching between:
  - audio or visual-based presentation;
  - graphical or text-based presentation.
- Limit the degree of freedom of choice of the learners. These choices are concerning the three elements mentioned and the freedom to switch between the instructional and encyclopedical mode.

The target of the project
In 1991 and the first half of 1992 the architecture of the database has been designed and the standard instruction routes are being developed for use as a prototype. During the remaining of 1992 the database will be completed. Early 1993 it will be used for research in real educational settings. The research will focus on the following questions:

- Do teachers use the modification tools?
- Which tools do they use and why?
- Are the results of their modifications according to their aims?
- Which types of tools are still missing to complete the teacher's facilities?

The presentation will focus on the elaboration of the principle of the teacher as the central manager of the database: The types of applications that are available (instruction / encyclopedia / presentation / learner-registration) and on the role of the teacher in using and modifying these applications for use in the classroom. Finally the design of the research activities related to this focussing will be introduced and discussed.

Abstract of:
Generic multimedia database: New didactics
Proposal submitted to the 1992 ECER theme: "Educational Technology: Instrumentation and Implementation" by R. Bestebreurtje, & P.W. Verhagen.

The goal of the current "Optical DataBase project (ODB)" is to construct a multimedia database for the following types of educational applications:

- Instruction (to individuals or in small groups). This is the major access to the database. It is organised by the construction of default "standard instruction route" for several target groups;
- Use of the database as a hypermedia encyclopedia;
- Classroom-presentations.

Central in the use of this database is the teacher as the manager of instruction. The database will contain several tools for the teacher to modify the database according to local needs. This modification implies:

- switch on/off specific topics;
- change the educational level of treatment;
- audio or visual-based presentation;
- graphical or text-based presentation;
- limit the degree of freedom of choice for the learners.

The presentation will focus on the evaluation questions about the tools for the teacher:

- Do teachers use them?
- Which tools do they use and why?
- Are the results of the modifications according to the aims?
- Which types of tools are still missing to complete the teacher's facilities?

Also the research design will be introduced.

THE USE OF VOICE SYNTHESIZER IN THE DISCOVERY OF THE WRITTEN LANGUAGE BY VERY YOUNG CHILDREN

Rachel Cohen, University Paris-Nord, France

The aim of this paper is to describe part of a research project carried out over a period of five years in France.
From previous researches, we know that young children aged 3 to 5 are able to : -discover the written language before age 6(1, 2), whatever their language, nationality, socio-economic background(3) ; learn a foreign language very early(4, 5) ; use a computer for learning purposes ; discover and acquire the written language if they are provided with an interactive creative piece of software(6).

The issues

The problem which concerns us to-day is to find out if the computer can make a major contribution to the teaching of a foreign language to children aged 3 to 6, both orally and in writing, thus combing all of the previous findings, and if the use of a voice synthesizer will affect the results.
A rapid word decoding ability and phonological awareness seem essential in the process of learning to read. Thus, the questions are :

- What would be the contribution of a speech output, as a tool, in such a specific form of learning
- Is the auditory feedback an essential part in the acquisition of reading or in the learning process in general ?

Target population

The school population we have been working with comes from different social environments. In this paper, we shall concentrate on a particular type of children : non-francophone children aged 3 to 6 attending pre-school in a suburb located in the North of Paris.
The district where this research project was undertaken consists almost entirely of non-francophone working-class families who are, for the most part, unskilled.
Many families do not speak french, and some parents are illiterate in their mother tongue. This explains why a large proportion of children start school careers with a definite handicap.
The challenge for us was great : can new technologies bring about a change in this situation if action is taken before school failure appears in the primary school ?

619

The problem

It is a complex one because it depends on several factors at once : psychological, linguistic, educational, social and technical.

The discovery of the French language as a foreign language, both written and oral, to non -francophone children below age 6, with the support of a computer equiped with a voice synthesizer is indeed an exceptional situation that upsets firmly-rooted beliefs.

It is generally believed that children below 6 cannot or should not learn a foreign language, cannot or should not learn how to read cannot or should not learn to read before learning to speak in a foreign language It is also discussed whether young children before age 6 are able to use a computer for learning purposes.

In this research another variable is added : the effect of voice input.

The program used

The computer is permanently in the classroom in a selfservice situation. The children go when they feel like it. The computer corner is a privileged place where interactions are most rich and numerous.

The computer " *which speaks*", is at first, an object of curiosity which in itself implies an element of questioning. The child is led to discover its possibilities by trial and error, then by a succession of precise actions. Very soon, the enthusisam engendered is such that access to the computer has to be regulated.

The program used " *Composition*" (6, 7) has been created by a group of educators, psychologists, computer specialists and technicians belonging to th "Learning Group", at the Centre Mondial Ressources Humaines in Paris, in 1985. The computer used is a BBC and the voice synthesizer is the system Ferma F 5000.

The program allows the child

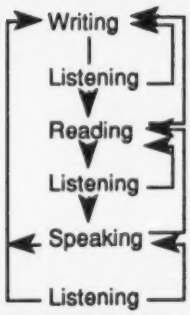
- to produce pictures by chosing words in a list offered in the program (global approach) or by typing them (analytical approach). When doing so the object appears on the screen. The child then can move it about at one's whim by the use of arrows. All the words (about 120) are pre-programmed, so that the child, when typing, can correct his own mistakes when he hears a "beep" .This rules out any risk of failure.Various topics are available : landscape, farm, beach, zoo, fairy tales, and these provide a great wealth of choices.
  Writing is, therefore, not an aim in itself, but a mean towards the realisation of a personal project which appeals to the imagination.
- to produce text in order to enable the child to write his own stories about the pictures he has created or any other story he chooses to write. In this case , the words are not obviously pre-programmed, but he has a number of means to correct himself without the help of the teacher.The texts become longer and longer, and more complex while the children get mastery of the tool and the language.
- to produce voice output : at any point, every letter, every word typed can be heard and listened to as often as desired. The child can also listen to the whole text if he wishes.

Findings

The problem for us was not to evaluate in the classical manner the peformances of the experimental group compared with a control group nor to ask ourselves : do these children perform better or not with such equipment ?

We were primarily concerned to observe what the children would do in this new learning situation, what would be their strategies to master this new tool, , to understand better the process of learning when sound produced by the child himself inteferes in the learning process and acquisition of language.
For all these reasons and to fit our purpose, we decided on qualitiative evaluation based on regular observations, case studies , analysis of productions, interviews of the teacher...
We shall concentrate on certain aspects that seem to us most important

Writing in order to read, reading in order to speak
That is indeed a strange statement which will shock many educators !
First of all , this new situation leads the children to rich verbal exchanges : comments, discussions ... Not only there may be several children in front of the computer, but they have a new playmate : the voice! The computer is thus personalized, it becomes part of the game, it is spoken to, or scolded if the computer does react the way the children would wish ! That is the starting point for motivation which is an important factor in the acquisition of language : the dialogue is based on the child's own experience, and state of knowledge enhanced by the voice synthesizer.
For the non-French speaking child, the image which appears on the screen makes clear the meaning of the written as well as the oral word . Contrary to the usual situations, it is the written word that comes first.
The fact that the letters are pronounced as they are typed makes the children aware of how the word is constructed and, intuitively they repeat the name of the letters as a game. Later they play at pronouncing the letter before the computer! When the word is pronounced, the child is then able to link the written word, its meaning (by the picture produced,) and its pronounciation.
Once the picture is finished, the adult will help the child to tell his own story, which becomes more and more elaborate. The teacher takes this opportunity to enlarge the vocabulary, and to correct the child's sentences.
A little text editor will then enable them to write the story on the computer and to print it. The child loves hearing his own story over and over again, especially when he has signed it which enables him to hear his own name.
Let us point out that the problem of "*writing*" is solved even if the child is very young and unable to master a pencil or pen, or if he is yet unable to form his letters by handwriting.
By a very interesting logical process, the children who have written their own texts, are able to read them (or guess most of what is written) whatever their age. They take pleasure in expressing themselves orally , with the help of the pictures they have created by means of written words.
Thus there is a constant interchange between the oral and the written language and the following diagram can be established :



This interaction between all processes of communication is dynamic, each phase interacting on all the others. At every stage, the level of communication becomes more elaborate. Let us notice though that *"listening"* is present at every level of the process.

For non-francophone learners, this process is even more interesting : it shows that writing and reading can develop simultanously with oral language. More than that , it shows that in this new learning situation, the written language comes first.

Important remarks :

- we cannot longer speak of *"reading" or writing" or* spelling as independant activites, but acquisition of language. No definite *progression* is followed : the children will use any word they need or know, however difficult it may be. *Spelling is* within reach of very young children : they acquire the words as they use them and the teacher does not have to ask them to repeat.
- the *knowledge* goes beyond what can be expected of young children. Example : when they write a text without punctuation, the voice synthesizer will read it without stops, and it makes no sense. The children react. The teacher then explains the need of punctuation and capital letters, Even children of 4 or 5 years of age are quite capable of mastering such notions.

Multiples sources of information and of self-correction

In a traditional situation of learning the teacher is the main source of information. In this case the situation is completely different. We have analysed 6 ways of finding information, some of which are linked with the learning situation as a whole and not specifically to the computer :

- *the computer program* which is self correcting - *the voice synthesizer* will allow the child to hear that a word is not written the correct way.- the children can go to *"the list"* if they are not sure of the word they wish to write.- *other children* : let us not forget that the computer is a excellent means of cooperative and group work. - *other sources*, such as books, dictionaries not forgetting those made by the children themselves.- *the adult :* if none of the other means seems efficient, then the children will go to the teacher.

The children's behaviour

Beyond the acquisition of knowledge, it seems to us that the construction of the children's personalities seems more important than the actual result of a picture or a text produced.

This learning situation encourages

- *Self- learning and self-organisation* : self-awareness develops and the effect of voice synthesizer is particular important in this respect.- *Self-esteem* : they are proud of their results whatever they may be. No frustration is felt if a mistake occurs because The voice feedback acts an positive reinforcement.- *Concentration and attention* : we have observed children below 5 years of age staying more than 45 minutes at the computer corner ; they have to be dragged away to go to recess or to go home ! - *Speed in the learning process* : the children work at a greater speed - *memory* : it seems that parallel auditory and visual feedback enhances the development of memory .- *Methods of work* : at the end of the year we were able to observe that children had developped their own method of work and that the previous acquisitions helped them to become more and more efficient. They also have learned to share, to cooperate between themselves. The voice makes the computer even more "sociable" than usual !- the most important of all is the *Peasure of working* : explosions of laugh and joy are heard around the computer, especially when the computer says" silly things" (mispelt words that sound "funny")

Conclusions

First of all let us say that children of all ages, whatever their nationalites, mother tongues, knowledge of French, are encouraged in their learning processes by this new learning situation . It seems clear that the voice input enhances the awareness and the speed of acquisition of learning, particuliarly in the type of education process we have described

Theoretically, this research made us think again about what we thought we knew concerning learning to read and the strategies of teaching and learning a foreign language.

References

Cohen R. (1992). L'apprentissage de la lecture, Paris, PUF 1977 ; 5°éd. traduit en espagnol, italien, portugais, bulgare.

Cohen R. (1982). Plaidoyer pour les apprentissages précoces, Paris, PUF.

Cohen R., & Gilabert H. (1986). Découverte et apprentissage de l'écrit avant six ans, Paris PUF ; 2° éd. 1987.

Cohen R. (1969.). Contribution à l'étude différentielle de l'efficacité des méthodes pédabogiques, thèse de Doctorat de 3°cycle, Université de Caen.

Cohen R. et al, Que sont-ils devenus ?, Les effets des apprentissages précoces, Paris, Revue de pédagogie, 88, P. 27_40.

Cohen R. et al, (1987). Les jeunes enfants, la découverte de lécrit et l'ordinateur, Paris, PUF.

Cohen R. et al, (1992.). Quand l'ordinateur parle, Paris PUF.

623

ELECTRONIC BOOKS AND THEIR POTENTIAL FOR DISTANCE LEARNING

Philip Barker, Interactive Systems Research Group, Teesside Polytechnic, United Kingdom

Abstract

The growing popularity and importance of electronic books opens up many possibilities for the dissemination of instructional materials. Electronic books are made possible because of the availability of low-cost, high-capacity storage facilities based upon the effective use of digital optical storage media - particularly, compact disc read-only- memory (CD-ROM). Because of their portability, robustness and pedagogic potential such books offer a powerful mechanism by which to implement distance learning programmes. This paper describes and discusses the special features and properties of electronic books that make them particularly suitable for the support of both open and distance learning activity. Two case studies are then briefly described.

Introduction

Education and training are intimately concerned with knowledge creation and knowledge transfer processes. Some years ago we analysed the role of conventional books in technical knowledge dissemination processes. We identified a number of important limitations of such books and suggested that some new form of book was needed in order to overcome the limitations of those that are printed on paper. We used the term 'electronic book' to describe a new form of book whose pages were composed from dynamic electronic information. We now use this term to describe information delivery systems that are capable of providing their users with access to pages of reactive electronic information with which they can interact. The pages of information which make up an electronic book are organised conceptually just like the pages of a conventional book.

A variety of different types of electronic book currently exist. They differ in many ways - such as the medium upon which they are published, the types of facility that they provide, the page structures that they employ and the media mix (text, pictures, sound, etc) that they use. Electronic books also vary according to the type of information that they contain, the way in which they are designed and the nature of the 'reader services' that they embed. Using these last three characteristics, it is possible to classify electronic books into nine basic classes. The classes that we propose are: text books; static picture books; moving picture books; talking books; multimedia books; hypermedia books; telemedia books; intelligent electronic

books; and vitual reality books. This paper deals primarily with telemedia books and electronic books that can be used as a basis for the creation of portable interactive learning environments.

Methodology

A considerable amount of research has been undertaken into the design and construction of electronic books and into their application for learning and training activities. We believe that such books have much to offer in the context of supporting distance learning programmes. They can form part of an integrated media approach to knowledge dissemination or they can be used in a stand-alone way. In any given situation the approach used will depend very much upon what is to be taught and the audience for which the pedagogic material is intended. In the remainder of this section of the paper we outline two case studies that describe some of the work that we are currently undertaking with respect to the application of electronic books to distance learning.

Case 1: Telemedia Books

A telemedia book is one which is intended to form part of an integrated media approach to distance learning in which the delivery stations (used for book access) are also capable of accessing the communications facilities available through wide area networks. The pedagogic strategies involved in using this type of book depend principally upon two basic paradigms: guided self-study and remote tele-tutoring. The telemedia book itself embeds all the instructional software and pedagogic material that is needed in order to create an interactive computer-based self-study environment. The knowledge corpus embedded on the CD-ROM contains all the resources that are involved in supporting the particular course of study being undertaken. Student performance is monitored, assessed and reported on by means of local 'intelligence' in the form of an embedded expert system. At times when it is appropriate students can participate in tele-tutoring activities with the tutors (either at the root organisation or elsewhere) to whom they have been assigned. They can also send email to each other (and to tutors) and join in computer conferencing activities. All the information (factual, procedural and executable) to implement these processes is contained within 'the covers' of the electronic books that are distributed to the students involved in a particular distance learning course.

Case 2: Portable Interactive Learning Environments

The idea of a Portable Interactive Learning Environment (PILE) is one which has considerable support from a large number of industrial and academic organisations within many countries. Fundamental to this approach to education and training is the concept of students and trainees 'carrying with them' the major part of their learning/training environment as they move from department to department within an organisation or from country to country within a multi-national corporation. The PILE project is therefore intended to examine the way in which libraries of self-contained electronic books can meet this requirement. The access stations are fitted with both a modem and a CD-ROM drive. These stations are used pedagogically in much the same way as those used to deliver telemedia books. They differ, however, in their portability.

Conclusions

The success of electronic books for the support of distance learning activitis depends critically upon four major factors. First, a suitable infra-sructure must exist in order to facilitate their authoring, publication and dissemination. Second, an appropriate publication medium (magnetic, electronic, or optical) must be available - this should be low-cost, high-capacity and very robust; for the work that is described in this paper digital optical storage in the form of compact disc read-only-memory has been used as the publication medium. Third, low-cost, portable interactive delivery stations must be made available; these will also need to be robust and easy-to-use. Finally, suitable learning paradigms must exist in order to facilitate both the development and the delivery of pedagogic material using the electronic book metaphor.

References

Barker, P.G. (1990.). Intelligent Electronic Books, Journal of Artificial Intelligence in Education, 2(1), 7-13.

Barker, P.G. (1991). Interactive Electronic Books, Interactive Multimedia, 2(1), 11-28.

Barker, P.G., (Ed). (1991). Electronic Books, Special Edition of Educational and Training Technology International, Volume 28, Issue 4, 269-368, Kogan Page, London.

Barker, P.G., (1992). Porable Interactive Learning Environments, paper presented at AETT '92 International Conference, University of York, 4th-6th April.

Barker, P.G., & Giller, S. (1991). Design Guidelines for Electronic Book Production, 179-182 in Proceedings of ICOMMET '91, International Conference on Multimedia in Education and Training, Tokyo, Japan, September 13-15.

Barker, P.G., & Giller, S. (1992). Electronic Books, 179-184 in Aspects of Educational Technology, Volume XXV: Developing and Measuring Competence, Edited by D. Saunders and P. Race, Kogan Page, London.

Barker, P.G., & Manji, K.A. (1988). New Books for Old, Programmed Learning and Educational Technology, 25(4), 310-313.

Giller, S. (1992.). Design Guideline for Electronic Book Production, MPhil Thesis, Interactive Systems Research Group, School of Computing and Mathematics, Teesside Polytechnic, Cleveland, UK.

Reimann, G. (1992.). Animation in Electronic Books, Outline PhD Proposal, Interactive Systems Research Group, School of Computing and Mathematics, Teesside Polytechnic, Cleveland, UK.

COURSEWARE ENGINEERING USING KNOWLEDGE-BASED PERFORMANCE MODELING TOOLS

Al-Noor Ladhani, & Italo De Diana, Department of Education, University of Twente, Enschede, The Netherlands

Introduction to the problem

Courseware Engineering, as a sub-discipline of Educational Technology, is concerned with the systematic design, development, use, and evaluation of courseware products. As a scientific field, Courseware Engineering studies the development, use, and evaluation of methodologies, techniques, and tools for product design, development, use, and evaluation. The emphasis in Courseware Engineering lies on increasing the productivity of courseware authors, and improving the quality of the courseware products that are made. In order to increase the productivity of an author, it is necessary to reduce the amount of time and effort involved in the development and implementation of courseware products.

A promising approach to realise this goal seems to be the development of authoring environments which are based on creating courseware product by composing these products from reusable components (De Diana & de Vries, 1990; Li & Merrill, 1990). These components are adjustable and exchangeable (also-called "pluggable" components").

It is assumed that adjustable courseware components are made by professional courseware authors, and are used by non-professional authors to compose courseware products, according to their specific desires.

Authoring environments for adaptable products, also called adaptation environments (De Diana & de Vries, 1990) or transactional shells (Li & Merrill, 1990), offer adaptation functions which enable the non-professional author or so-called adaptation agents to inspect, add, remove, and adjust exchangeable components of a courseware product. This approach could effectively support large scale use of courseware products.

An essential factor for this approach to become effective is the support given to adaptation agents while composing and adjusting a courseware product. Merely providing adaptation functions will probably not suffice to enable agents to easily construct courseware products according to their preferences, and of sufficient quality in respect to the desired learning goals. We expect that it is necessary to expand adaptation environments with functions that provide adaptation agents with guidance for product component selection, and their adjustment.

627

In this respect, we suggest to expand adaptation environments with performance modeling tools (PMTs). These are software modules capable of predicting specific performance of courseware products when provided with information about characteristics of these products (Ladhani & De Diana, in press). PMTs enable adaptation agents to inspect the likely effects of design choices on aspects of courseware product performance. For example, the adaptation agent could use a "learning-time" performance modeling tool to exam the effects of a specific learning strategy on expected learning time.
In PMTs, product characteristics are viewed as "structural variables", as defined by Card, Moran & Newell (1983). Examples of some instructional structural variables are instructional goals, student characteristics, organisational strategies, delivery strategies, and management strategies (Reigeluth & Merrill, 1979). Product performance can be indicated by performance variables as defined by Card, Moran & Newell (1983). Examples of some instructional performance variables are accuracy, speed, transfer, retention, learning time, and appeal (Reigeluth & Merrill, 1979).
The use of performance modeling tools is likely to improve the productivity of courseware authors and the quality of developed courseware products. By using PMTs, the author is encouraged to reason in terms of performance of courseware products instead of merely considering, and being driven by detail development and implementation matters. Card, Moran, and Newell (1983) show that this process of attention focusing has an important positive influence on the productivity of a designer and on the quality of the resulting design. We expect that research on PMTs for courseware development can provide a significant contribution to the field of Courseware Engineering.

Knowledge representation in PMTS
Performance models can be roughly categorised into experimental, database and symbolic models (Card, Moran & Newell, 1983). In experimental models, performance variables are obtained by performing experiments and measuring the outcomes. Database models are sophisticated lookup-tables which store pre-measured or pre-calculated data. More interesting however are symbolic models that allow to establish performance outcomes that are not pre-measured or pre-calculated. We classify symbolic models as mathematical models and as knowledge-based models. Mathematical models contain mathematical equations representing a calculational relation between structural variables and one or more performance variables. Knowledge-based models contain a knowledge base in which the knowledge to deduce or induce performance prediction from structural variables is explicitly represented in some formalism such as rules or frames.
The knowledge-based model seems most suitable to represent the knowledge for predicting instructional performance variables. Instructional knowledge that is explicated from human experts and literature mainly consists of if-then type of rules. A suitable way to represent this knowledge is in knowledge-based models. We will refer to this type of PMTs as knowledge-based PMTs.

Knowledge-based PMTS
The following elements of a PMT, and its relationship to the other components in courseware authoring can be described. The input of a PMT consists of values of

structural variables. The values of structural variables are extracted from the composition and instantiation of an adapted product. The knowledge that is necessary to perform this extraction is called translation knowledge. The values of structural variables are stored in a so-called semantic model of the courseware product. Next, the PMT is used to predict the values of one or more performance variables. The knowledge that is necessary to make the prediction is called interpretation knowledge. This outcome is stored in a so-called pragmatic model of the courseware product. An appropriate human-computer interface is used to present the performance information to the adaptation agent.

An example of a knowledge-based PMTS

We have planned to prototype a knowledge-based PMT that is capable of predicting the expected time for a learner to complete an interaction with a courseware product. This small but concise example should illustrate the functions and the machinery of a knowledge-based PMT.

Conclusions

In order to be able to "engineer" the quality of courseware products, knowledge-based performance modeling tools are to be developed that can provide explicit guidance for the selection and valuation of design variables. Such tools should focus the attention of author on important structural and performance variables while making design choices. Such focusing is expected to increase the productivity of the design process and the quality of the resulting products.

References

Card, S. K., Moran, T. P., & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, New Jersey: Lawrence Erlbaum.

De Diana, I. P. F., & de Vries, S. A. (1990). Adaptable educational courseware: An antidote to several portability problems. Journal of Research on Computing in Education, 23(2), 225-242.

Ladhani, A. N., & De Diana, I. P. F. (in press). Computer aided courseware engineering and performance modeling tools. In I. P. F. De Diana & P. van Schaik (Eds.), Lecture notes in computer science: Aspects of courseware engineering. Berlin: Springer-Verlag.

Li, Z., & Merrill, M. D. (1990). Transaction shells: A new approach to courseware authoring. Journal of Research on Computing in Education, 23(1), 72-86.

Reigeluth, C. M., & Merrill, M. D. (1979). Classes of instructional variables. Educational Technology, 19(3), 5-24.

DESIGNING "SHARE": AN ENVIRONMENT FOR CLASSROOM-BASED TEACHING AND LEARNING

Zhiting Zhu, & Italo De Diana, Department of Education, Enschede, University of Twente

From Classroom Information Systems to SHARE

The development and employment of Classroom Information Systems (CIS) has constituted an important line of computer use in instructional processes in some oriental countries, such as in China, Japan, and South-Korea. Typically, such systems are used for learning foreign languages, but other types of applications, such as mathematics, can be pointed out too. The spectrum of this type of systems, which originated from CMI applications, is featured by central control of audio/video media, group-based presentation of questions/test items, real-time gathering of students' responses to each item, and dynamic report and analysis of response data. It is characteristic of these systems that they usually are driven by a single PC, that is, a single PC is used to serve up to 64 learners (Zhang & De Diana, 1990). The authors of the present contribution believe that the functionality of such kind of systems could be upgraded to become Classroom-Based Teaching and Learning (CBTL) systems by augmenting the authoring functionality and the delivery capacity, integrated with improved hardware design. In this paper the design of such a type of system, named "SHARE", is described, with emphasis on its design philosophy, authoring machinery, and some options for product portability.

About CBTL

A newly designed CBTL system, which has been named 'SHARE', is described. SHARE is to serve up to 64 learners with one single PC, and yet it is to offer some multi-media facilities, options for developing and changing communication structures, and using various sources of learning materials. SHARE is to build on available hardware and software technology, but it is to extend some available concepts in order to allow so many learners to use shared, and rather simple, facilities. Several reasons can be given for the development of SHARE. An important reason is cost. In a country like China, computer hardware is expensive and scarce. Another one is educational philosophy. Whereas Western countries strongly stress individual learning and personalized instruction, some Asian countries use group-based approaches to teaching and learning. SHARE is in line with this philosophy. Learners and teacher cooperate in a group-based teaching and learning environment.

Some features of SHARE

The SHARE system provides both authoring facilities and a delivery mechanism, serving for the production of courseware and the support of instructional transactions. Four different modes of CBL are to be developed: tutorial, (simulation-based) demonstration, drill and practice, and testing.

Different from most existing CBI systems, SHARE includes the classroom teacher as a crucial part of the system. He might, to some extent, play the role of a dynamic instructional strategy, an instructional information source, or a learning monitor, dependent on the difference of instructional phases and instructional strategies currently used. Some distinctive features held by the SHARE sytem are described such as:

- Support of close-circled instructional processes and open-circled instructional processes, in the former situation the delivery of instruction by the computer take place automatically whereas in the latter situation instructional presentations are interleavedly given by the computer and the teacher.
- The immediate availability of analytical reports for the teacher, as a consequence of real-time gathering and analysis of students' responses.
- Dynamic grouping of students in the light of the data gathered during learning their course.
- Multi-way communications among different agents involved in the learning environment: thé computer, the classroom teacher, student individuals, and student sub-groups.
- The trade-off between group-based pacing and individual-oriented presentation. The latter is accomplished by the hyper-screen display, an expanded concept of hypertext, which schedules different hyper-frames pertaining to a single cluster of concepts in according to students' choices and produces a different display on each student's desk-top screen.

Some Portability Considerations

The design of the SHARE authoring environment involves an integration and redesign of the EDUC-USE system (De Diana, 1988), a tutorial-oriented authoring system, and the MCOGEN system (Zhu, Wu, & Jiang, 1989), a hybrid authoring instrument intended for use in Chinese education. In consideration of the potential of international use of the SHARE system, several cross-cultural educational differences between China and Western countries, which could been seen as significant factors affecting the portability of SHARE, particularly the authoring instrument, are under our design consideration. These factors include:

- The communication medium: an icon-based language (Chinese) versus symbol-based languages;
- The instructional approach: group-based versus individualized;
- The curricular structure: a problem-driven versus an explorative approach.

The portability of an authoring instrument can be described as the degree to which software can be transferred from one environment of use to another. In other words, the portability problem is caused by the differences between the source environment (for which it was developed) and the destination environment (in which it is to be used).

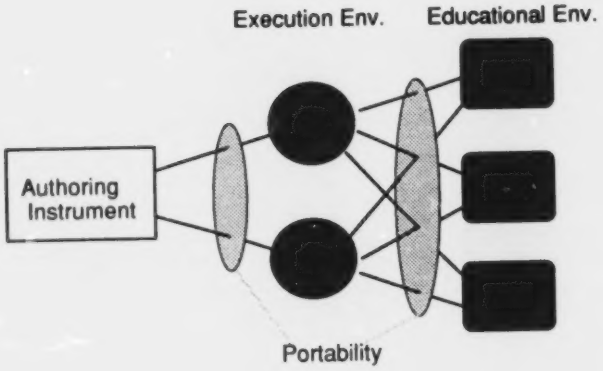


Figure 1
Authoring instrument, environments, and portability

The use environment comprises two layers: the execution environment in which the software is to be executed, and the educational environment in which the software is applied (Figure 1). The portability problem caused by the differences between various execution environments could be solved using techniques from software engineering, such as abstraction, modularity, and reuse. In other words, this type of portability problem imposes a need for a flexible software architecture. Whereas the portability problem caused by the differences between educational environments necessitates an authoring instrument to have such functional flexibility that it can produce courseware products adaptable to different educational needs. Thus portability considerations urge design requirements for both a flexible software architecture and a flexible authoring functionality.

The authors of this paper believe that the concept of a flexible authoring system could be a useful solution to the portability problem as described above. A flexible authoring system refers to an authoring system that possesses a sufficient level of functional flexibility to meet different users' style of use and application needs. As suggested by Zhu and De Diana (1992), the flexibility of an authoring tool can be described on eight dimensions: authoring mode, applicability, cultural interface, authoring expertise, domain knowledge representation, course organization and management, instructional strategy, and information creation and presentation. The design of the SHARE authoring system receives attention in respect to flexibility on the following dimensions:

- Cultural interface
- Information creation and display
- Instructional strategy
- Learning management

Conclusions

The design of SHARE suggests that it is valuable to take educational and cultural differences between the Eastern and the Western world, or among different countries in a more general sense, into account when designing and developing computer-based learning systems. Portable authoring and learning environments could be expected to play a role in bridging educational and cultural differences and thus promote the scale of development and use of Computer Based Learning.

References

De Diana, I. (1988). Het EDUC systeem : Aspekten van een methodologie, ontwikkelingsmethode en instrumentatie voor tutorieel CQQ (The EDUC system. Aspects of a methodology, development method, and instrumentation for tutorial CBL). Ph. D. Thesis. Enschede: University of Twente.

Zhang J., & De Diana, I. (1990). Collaborative computer-based learning in China: a system for classroom-based learning and teaching. In D.H. Norrie, H.W. Six (Eds.). Computer Assisted Learning, 3rd International Conference, ICCAL '90, Hagen FRG, June 1990 (pp. 442-452). Lecture Notes in Computer Science 438, G. Goos, J. Hartmanis (Eds.). Springer : Berlin.

Zhu, Z., Wu, G., & Jian, Y. (1990). MCOGEN: a model-based courseware generation system. Proceedings of 4th National Conference of the Chinese Association of Computer-Based Education, 2. Tianjin (China).

Zhu, Z., & De Diana, I. (1992). Flexible authoring environments. In I. De Diana & P. van Schaik (Eds.). Aspects of courseware engineering: Educational instrumentation perspectives (Lecture notes in computer science) In press. Berlin: Springer Verlag.

INTERACTIVE AUDIO LEARNING ENVIRONMENT

Ji-Ping Zhang, Department of Educational Information Technology, East China Normal University (ECNU), Shanghai, China

Abstract

This paper offers a description of a computer-based audio learning system. Based upon a PC computer and audio, this system is intended to support the learner to study in an interactive and effective way, especially for learning languages. During study, the learners can not only see and hear from the computer, but speak with a microphone. This system is mainly composed of a CP-286 computer with harddisk, an audio card and a stereo recorder. It is possible for the audio card to Direct Memory Access (DMA) with harddisk of computer and input the voice to harddisk using a microphone or recorder. The voice quality is nature and good.

Introduction

Nowadays personal computers with userfriendly software are to be found everywhere in the school as wall as in the home or office. But that does not mean that all problems have been solved or that no new problems have been arised and that all available courseware meets acceptable standards.

One of the problems for CBL system, the output was, and mainly is of the visual kind. As everyone knows, it is not enough to just have the visual for education, but the audio also is important. During the eighties, some countries have developed out several ways for adding audio to the CBL system: such as CD-I, voice synthesizer, digital sample and tape recorder. The CD-I technology has a high quality sound and quick search function, but it is expensive for an individual CBL system at present. Especially it is very difficult for people to make the CD-disk by themselves. For the voice synthesizer, the primary factors limiting its use as an audio display for CBL system are the c of memory for storing the voice. For instance, one second of voice usually needs over eight thousand bytes of memory. Which way is more suitable for us to add audio to personal CBL system in our country today? The answer is the tape recorder, because we think the tape recorder at least has following advantages: First, it is the cheapest to produce the sound. Second, it has a good enough sound quality to satisfy the requirements of education. Finally, it is easy for each person to record the sound on the tape by themselves. Of course, the restrictions of the ordinary tape recorder are that it plays sequentially in one direction and is unable to do quick and free search of segments on the tape. If we can overcome these restrictions by some methods, the

tape recorder will become a low cost and effective audio media for CBL system. In order to find this solution, we began our study a year ago. Our purpose was to solve the problem of the quick and free search. After repeated experiments using several available recorders, we have developed a prototype of desired recorder which Can be controlled by computer and connected with an audio card.

Method and principle

The interactive audio is mainly composed of a stereo recorder and an Interface Control Board (ICB). The ICB serves to exchange the massage between the computer and recorder and to control all operations of the recorder. The ICB mainly consists of three parts: Bus buffer, Voice signal detect, Segment search and counter. The bus buffer is responsible for the bidirection signal transfers between the computer and ICB. It consists of the status register, tristate 110 logic gates and address decoder. The function of voice signal detect is mainly to detect the voice signal which is coming from the recorder and then to convert it into a logic lever that will be sent to the status register. The computer will know via the status, when a voice segment is ending. In fact, there are a lot of segments on serves to computer screen. The period of silence (3 to 5 seconds) between two segments is called as an event. The event coming means that the present segment has ended. There are two tracks on the tape. We can record the different contents on each track. For instance, L track is English and R track is Chinese or one of them is music. At any moment, it is possible to switch on one track or turn two tracks together. This is why we select a stereo recorder. The search part for segments serves to perform the function of segment search in free and quick on the tape. This part mainly consists of a Quick-Music-Search (QMS 7512) chip and a special counter. As we know, a ~MS 7512 chip can just offer to search three segments that is not to satisfy the needs of instruction procedure, because a courseware normally is composed of a lot of segments. In order to solve this problem, we used a special counter and developed a program which can help the ~MS chip to search the segment number with unlimited. The counter and program also can make the QMS chip to know how many segments should be jumped and which is the present segment. In fact this recorder has two working modes: On-line and Off-line. In the off-line mode, it is like a standard recorder and each operation can be done by touch the function keys. When it is working On-line, all operations will based on the control command from the computer. There are three kinds of commands: Control Commands, Detect Commands and Combined Commands. It is easy to interface this audio to an available CBL courseware. The things that we should do are to record a tape and insert some commands into courseware. During the course, the event will be detected. The event coming means that the present segment has ended and you can input your voice or repeat just segment, or go to next part.

A courseware example

This audio as a media is suitable use for most CBL system. If you want to connect this audio with your courseware, only a little program has to be added. Here we will present an example of audio CBL system as an inspiration for you. It is a Chinese language CBL course. Formerly, It was a courseware only could offer the student to learn the grammar, reading and vocabularies on the screen. After using

the audio, the function of this system has been improved greatly in which especially the function of the audio can enhance the student motivation for learning and drill students in listening comprehension. This course mainly has three parts: the situation dialog which can help the student to learn some words and sentences according to each kind of daily situation: the sentence learning which shows the most useful sentences of Chinese; the practice part which drills and tests the student ability for Chinese language. It is possible for each part to call the audio drive procedure and output the synchronized voice to the learner. The audio drive procedure is a short program and can be called and linked by any courseware which is developed in Turbo pascal 5. 0. Of course, it is necessary for you to record a related tape for the course. The using of the audio makes the courseware program to become shorter, because most of display can be voice instead. Following Chinese language, we also have developed some other language courses with audio, such as the minorities of Chinese, Russian, English and Japanes. they were all successful to link with the audio. When running a course, the student can see and hear a segment for several times and jump backward or forward from the present segment to another one in which the audio will always synchronize with the screen. At any time, it is possible for student to return the menu of the course to make a new choice.

References

Zhang, J.-P. (Agu. 1 986) Method and Technology of Voice on CAI system. MCSE'86 international conference. Tokyo Japan.

F. Jaspers, & Zhang, J.-P. (July, 1990). Interactive Audio for CAL. Journal of Educational Technology System. U. S. A..

THE VALUATION BY STUDENTS OF THE USE OF COMPUTERS IN EDUCATION

B.G. Doornekamp, Centre for Applied Research on Education (OCTO), University of Twente, Enschede, The Netherlands

Abstract

Two pilot schools for non-vocational secondary education in Enschede are taking part in the Experimental Schools Project. One of the research questions is the valuation by the students of the use of computers in education. The valuation is influenced by factors which deal with the perception of the student, the use of computers in teaching-learning situations, the circumstances in which the computers are used and background information on the students.

The research question is investigated by means of two identical instruments: a paper-and-pencil questionnaire and a computerized questionnaire. Each questionnaire will be administered to approx. 450 students. Prototypes of each version were administered to a small groep of students.

At the Conference the results of this study will be presented.

Introduction

In 1987 the Dutch Minister of Education and Science started the Experimental Schools Project (ESP). Two pilot schools for non-vocational secondary education in Enschede are taking part on this project. One school comprises of about 850 students and 65 teachers, the other school comprises about 1,400 students and 90 teachers.

The schools are equipped with 30 computers each, spread over computer classrooms and some stand-alone sets used for demonstration in the teachers' own classroom.

The schools receive funding for 25 hours a week of released time for teachers participating in the project and 10 hours a week of released time for two computer co-ordinators. Furthermore, the participating schools receive a budget for the purchase of courseware and the organization of in-service training.

A team of researchers of the University of Twente is conducting a programme of research on the implementation process of information technology in these schools.

This paper deals with COMPRO (Computer in Pilot-schools), one of the research projects of the ESP. Since the start of this project the research questions have focussed mainly on the teacher, but in the last stage of the project attention is also paid to the students (Van Diggele & Carleer, 1990).

Research question

Many teachers of several subject areas have been using the computers in the computer classroom since the ESP started. Some students have used the computer very frequently. This depends on the teachers who teach them.
When the students have lessons in the computer classroom, they are very enthusiastic. They like these lessons better than the ordinary lessons. But one can imagine that after many lessons in the computer classroom the students will not be so enthusiastic as before. The novelty-effect will have disappeared. Or will the students still be enthusiastic?
The research question is whether the students still like to use the computers: how do the students of non-vocational secondary education value the use of computers in education and which factors influence this valuation of the students? (Doornekamp & Carleer, 1992).
The following aspects are of importance in this research question:
the perception of the student;

- the use of computers in teaching-learning situations;
- the factors that are related to this use of computers;
- the background factors of the students.

The students are the most important target group of the use of computers in education. Their opinions on computers and courseware are extremely relevant. A side-factor is the rather unique situation at the pilot schools. During several years the use of the computer classroom has been recorded. Information is now available on the past history of these students with regard to their use of computers.
The research question focuses on teaching-learning situations. In those situations the computer is used to achieve a goal. In doing so, the use of the computer as teaching method is distinguished from other teaching methods.
The factors that influence the valuation can be divided into two groups:

- Factors which deal with the use of the computer itself. The factors are qualities by which courseware packages can be distinguished from each other. The students will be questioned how they value these qualities of a particular courseware package. Examples of these qualities are: interactivity, individualization and practice.
- Factors which can be carried back to the circumstances in which the computers are used. During a lesson the students work alone or in couples while the teacher helps them on request. Examples of these factors are: alone or in couples, courseware, guidance by the teacher and availability of equipment).

A number of background factors play a part in the interpretation of the differences in the valuation of certain group of students. The factors which count in this study are (among others): school, grade, gender, age and previous experiences with computers.

Method

The research question is investigated by means of a questionnaire. The questionnaire consists of two parts. The questions of the first part refer to the general valuation of the use of computers at school. In the second part the student has to value two out of five courseware packages. For each grade (of each school) the five best used packages were selected. A student only needed to give his

valuation for two packages he has used in the present school-year. From a try-out we learned that when the student have to do three or more packages, they loose concentration. Therefore the number has been limited to two.
For this study two versions of this questionnaire were developed: a paper-and-pencil questionnaire and a computerized questionnaire. Both questionnaires contain the same questions, only the way the questionnaire is administered differs. It is not the aim of this study to compare both instruments, only to gather some experience with a computerized questionnaire.
Usually a paper-and-pencil questionnaire is used to collect data from a large group of respondents. It happens very frequently that the respondents do not fill out the questionnaire correctly: questions are skipped or questions that do not apply are answered. After the questionnaire's administration, the data have to be entered into the computer before they can be analyzed.
By using a computer the respondents can be guided through the questionnaire. They only need to answer those question that apply to their situation. The remaining questions are not shown. The computer also makes it possible to record the answers of the respondents. After administration, the data are ready to be analyzed.
First, the paper-and-pencil questionnaire was developed. The content is based on the factors mentioned above. Some questions regarding background information on the students were also entered into the questionnaire. The background information regards, among others: gender, age, grade, computer at home, (previous) experience with computers, choice of education after secondary education and career planning.
In spring of 1991 the prototype version of the questionnaire was administered to a small group of students (n = 31) of both schools. After this questionnaire had been adjusted, the development of the computerized version started. For this purpose the authoring system TAIGA was used. The prototype version of the computerized questionnaire was administered to 20 students of both schools in December 1991. For this study an AT-computer (Olivetti) with a colour screen was used.

Design
From each grade (grade 7 up to and including grade 12) three classes (lower, medium and upper level) will be selected. Each selected class (approx. 25 students) will be split up into two equal groups. Girls and boys will be spread equally over both groups. One group will fill out the paper-and-pencil questionnaire and the other the computerized questionnaire.
Per pilot-school approx. 450 students will participate in this study (approx. 225 students fill out the paper-and-pencil questionnaire and approx. 225 students the computerized version).
The questionnaires will be administered at the beginning of 1992 (January, February and March).

Results
The data will be analyzed in the period between the administration of the questionnaires and the conference (June 1992) After the analysis the final paper will be written.

639

At the conference the results of this study will be presented (for the first time). An answer will be given to the research question mentioned above. Special attention will be paid to those students who have visited the computer classroom very frequently during consecutive years.
In the discussion attention will be paid to the results of this study and (if possible) some recommendations will be made for the schools of secondary education which are about to implement computers in their education. Some experiences with computerized questioning will also be discussed.

References

Diggele, J.B.H. van, & Carleer, G.J., (1990). Computergebruik en courseware-selectie door docenten van de proefscholen. Voortgangsverslag COMPRO-project. Serie "PRONTO-rapporten" nr. 7. Enschede: Universiteit Twente - OCTO.

Doornekamp, B.G., & Carleer, G.J., (1992). Courseware in het curriculum: op weg naar integratie. Voortgangsrapport schooljaar 1990/1991. Serie "PRONTO-rapporten" nr. 18. Enschede: Universiteit Twente - OCTO.

CLASSROOM CHANGE: IS THERE A STEERING EFFECT OF COMPUTER TECHNOLOGY?

John Olson, & Larry Miller, Faculty of Education, Queen's University Kingston, Ontario, Canada

It is often said that computers will change how teachers teach. Amarel (1983), for example, suggests that computers will steer schools towards an emphasis on "thinking skills" and a shift in the teacher's role from controlling person to facilitator of autonomy. he very basis of the work of the teacher is thus influenced: the curriculum becomes more generic and the teacher less influential. Such predicting of major steering effects for computers is common.

The history of innovation in education teaches us to be cautious about predictions associated with new technologies. Often the emergence of these new technologies merely provides a context in which to scourge the school for not lining up with visions of reformers -- in our example the familiar demands of the psychologically minded for a scientific technique that circumvents traditional forms of curriculum and teaching.

Beyond the caution of history, what can we learn about the supposed steering effects of computer based learning on the curriculum and its enactment? In this paper we use data from a number of case studies from two research projects to argue that existing innovative classroom practice is amplified but not steered by computer resources, and that the existence of the innovative practice has much less to do with the advent of technology than with the teacher's pre-existing conception of practice. Although the technologically minded see computers as instruments for a new classroom laboratory, our results suggest that pre-existing pedagogical intentions steer the interaction between approach to the subject and computer resources. For example, teachers eager to stimulate inquiry do so before they adopt computers.

The first set of studies come from the Schools, Computers and Learning Project study of how an elementary school used ample computer and human resources over a three year period (Miller and Blackstock, 1989). The K-6 (ages 5 to 12) school involved in the study was a fertile site, with a ratio of approximately one computer for every seven students that were placed in a computer laboratory as well as individual classrooms. Complementing the networked computers were peripherals such as midi-interfaces, robotics kits, printers, and modems. Although the school involved in the study was rich in technology, most of the staff had little or no experience using computers, a situation different from typical lighthouse schools where teachers frequently are selected specifically because of their expertise with technology.

Two apparently contradictory results emanated from this investigation. Word processing, considered by many educators to be one of the most productive tools associated with technology, and an application praised as a potential change agent in terms of fostering how writing should be taught in schools, was used infrequently in the Grade 1 classroom. One reason for this lack of use lies in an inherent inadequacy in the painting/word processing program used by the children, but a more important explanation may be found in the teacher's usual teaching practices and routines as well as her vision of young writers' capabilities. A process approach to writing, such as one advocated by Graves (1983), was not evident in the classroom. Instead, children frequently were given dictation based on group stories or were asked to write sentences to accompany pictures. The teacher's view of children's capabilities as writers did not include self-selecting of topics, creating natural spellings, or revising composition efforts.

On the other hand, the teacher used data bases as an integral aspect of teaching over a two year period. Data bases appear to be a complex use of technology, especially when compared to word processing. This seeming contradiction may be explained when one looks at the Grade 1 teacher's past practices as well as her vision of children's capabilities. Each year, the teacher begins the term by carrying out a thematic unit entitled All About Me. Other themes follow. One aspect of this initial unit engages the students in gathering data about themselves. It includes such information as physical features, personal likes and dislikes, and house locations. The information is placed in a matrix, using chart paper, with children's names in the vertical position and the data registered horizontally. In essence, the children create a data base; thus, when the teacher was exposed to a computer program data base she embraced this technology. Over the course of the SCL project, she used the data base on numerous occasions, teaching the children to carry out the necessary research, enter the data, and retrieve it by playing a game built into the program.

The second set of cases comes from a year long study of teachers given additional computer resources in order to pursue their classroom possibilities (Olson, 1988). In the study a common thread amongst teachers asking for access to computers is the pre-existence of innovation practice and the desire to augment this by means of computers. The technology becomes an instrument to develop the practice and a means to express membership in the avant-garde. The computer thus has instrumental and expressive functions in classroom life. Rather than the computer steering their practice, the teachers entered into a dialogue with their teaching based on their experience with new technology.

The first teacher, Mrs. Everett, used the computer as part of her writing process instruction. Having to pay close attention to her students as they explore the computer medium and as they draw her into this process has led her to ask interesting and important questions about what she is trying to do. Granted she is not pursuing these questions in a systematic way, but she could have, had the pilot project design been more attentive to the teacher education possibilities of the process. The questions she raises are fundamental. They have to do with the different ways in which children learn, with fairness in the way resources are made available to students, with her own capacity to manage new ways of learning. At issue is the impact of her teaching on students, and what student behavior means. Why do some students persist and others not? Why are some students hindered in

writing by using the computer? Such observations are made against the grain of recommended practice--computers are supposed to help. Perhaps there are difficulties in the very writing method she has adopted.

Clearly there are difficulties in her use of the computer. Is it fair? Only some students are favored. Is it effective? Some students are disappointed--some do not write as well using the computer. She is burdened by considerable work. These difficulties do not seem to trouble her. There is an odd inattention to them. Why is this so? The computer seems to offer her a way of expressing her interest in avant-garde approaches to her subject. In her school board, the computer "speaks loudly" in a way that this or that pedagogical theory in her subject does not. Her avant-gardism appears to be directed outward beyond the school. Meanwhile there are problems in her chosen manner of expression--instrumental problems which are likely to persist. How long will it be before those problems make it difficult for her to use the computer in this strongly expressive way? What will she learn from the experience? These things remain to be seen.

Mr. Devon used computers to process geographical data, and, like Mrs. Everett, made the move to the computer from an existing approach which might itself be called innovative or avant-garde. Rather than having to meet some expectation about the "level of use" of the innovation, they have become the very architects of what the innovation is to be through their analysis of the experience. That experience, when reflected upon, becomes the basis for new practice, and more so in Mr. Devon's case who has responded in a variety of ways to his experience. He has developed appropriate seat work, for example, to accompany computer use. Without this, he said, he could not use the computers. He spoke about getting used to teaching through multi-group activity and how this was different from whole-class teaching in terms of the amount of material it was possible to cover, the classroom climate and demands on the teacher. Furthermore he used student "experts" to help him manage computer activity. He encouraged students to talk to each other at the computer and he enjoyed one-on-one contact with them in the multi-activity setting. Student enthusiasm for computer-based work encouraged him to make the major effort required to do it.

His experience has left him wondering about the usefulness of "standard" geography and "standard" teaching. No clear resolution of these issues emerges in what he says, but we sense his concern to be an important, helpful person to his students and to convey to them through avant-garde means his enthusiasm for his subject: for the thought his subject can provoke. There are doubts, however. Is all the time plotting the base map worth it? Is he wasting time doing that? Should he be dropping traditional topics like the Niagara fruit belt? Is the extra time he has to devote to his computer-based approach worth it?

These case studies comprise four teachers, and represent several grade levels and areas of the curriculum. The specific areas of technology studied include word processing, data bases, and simulations. The focal point of our paper will be an examination of the use of technology through the filter of teachers' existing beliefs, practices, and classroom routines. Both sets of cases suggest that the teachers' view of the subject and of pedagogy is more important than either the nature or extent of the computer resources made available in determining the outcomes when these resources are developed. Existing practice reproduces and amplifies through but is not steered by computers.

References

Amarel, M. (1983). Classrooms and computers as instructional settings. Theory into practice. 22, 260-266.

Graves, D. (1983). Writing: Teachers and children at work. Portsmouth, NH: Heinemann.

Miller, L., & Blackstock, J. (1989). Yes, but are they learning anything? Toronto: Queen's Printer for Ontario.

Olson, J. (1988). Schoolworlds/Microworlds: Computers and the culture of the classroom. Oxford: Pergamon.

TWO EXAMPLES OF SIMULATION ENVIRONMENTS FOR OPERATOR TRAINING BASED ON PARALLEL INSTRUCTION

F. B. M. Min., P.G. van Schaick Zillesen, H.B. Reimerink, M. R. Gmelich Meijling: Department of Education; G. de Niet, Department of Chemistry, University of Twente, Enschede; & F.J. van der Gaag, AKZO Salt and Basics Chemicals, Hengelo, The Netherlands

Summary
At the University of Twente, the Department of Education (Toegepaste Onderwijskunde) a design system was developed in the period 1984 to 1988 with which a trained courseware designer can make any computer simulation program in the fields of physics, biology, medical sciences, economy and technical installations.
The courseware, built with this system, is designed accoording the 'MacTHESIS philosophy' and based on a theory of parallel instruction.
Two programs are showed:

Computer simulation program: BRINE PURIFICATION
A simulated factory for brine purification, built for the chemical company AKZO Hengelo. We show a version with for three monitors and a 'parallel instruction system' in a 'multi-processing' environment on one Macintosh computer;

Educational aspects
designed accoording the 'MacTHESIS philosophy'
- multi-windowing
- coached problem solving environment
- parallel instruction and information system

Technical aspects
- 3 monitors
- 1 Macintosh computer
- model-driven computer simulation program
- evnt-driven, multi-windowing programs
- multi-tasking environment ('MultiFinder')
- parameter transfer by file communication

Computer simulation program: COAGULATION
An interactive 'model-driven', 'desk-top-video' program, named COAGULATION (STREMMEN) a simulation of the casein micell coagulation (enzym kinetics) in the cheese process.
The prototype we show, is experimental and is built for an optical database project ('ODB'-project) around a cheese factory.

Educational aspects
- designed accoording the 'MacTHESIS philosophy'
- interactive learning environment with video
- video fragments as feedback to the trainee

Technical aspects
- interactive model-driven 'desk-top'-video
- videocard: 'MoonRacker'

ORGANISATION OF COMPUTERS APPLICATIONS IN DUTCH SECONDARY EDUCATION

Th.P.J. Ris, Faculteit Educatieve Opeidingen, Hogeschool Rotterdam E.O., Rotterdam, The Netherlands

In 1984 the Dutch Government gave the green light for the so-called 'NIVO' project (New Information Technology Secondary Education). This project had several objects. One of these objects was to provide hardware to the participating schools. All of a sudden Dutch schools were provided with numerous computers that could be used in the different disciplines. In practice this caused and still causes serious organizational problems. For instance, who is responsible for purchasing supplies for the available computers, and when should these supplies be bought? To these and many other questions the present organization does not provide satisfactory answers.

The main subjects of this paper will be the feasibility and particularly the necessity of a sound organization structure within Dutch Education as regards Systems Engineering. At this stage it seems appropriate to further define this discipline. Systems Engineering is a discipline with among other things the unsaid task to integrate the application of computers in the existing structure of Education. The organization of Dutch Secondary Education is unable to carry out this task satisfactorily. The structure of this organization is not geared to this. In the past, this organization may have functioned efficiently. More than in other (private) organizations, in Dutch Education the substance and the consequential organization are kept strictly separated. Teachers, who are responsible for the substance, often adopt an aloof attitude in matters concerning the overall organization of Education. With the rise of the discipline Systems Engineering this controversy is often a source of inefficiency. For, this discipline often stands in the centre of the various organizational lines. All disciplines use the facilities of Systems Engineering. This inefficiency does not only manifest itself in the factor labor (system management / automation coordinator) but especially in the means of production (educational tools). The factor labor not playing a major part (yet) is mainly due to the fact that usually the costs thereof are systematically excluded from the Dutch Education budget. This fact seriously hampers the analysis required to arrive at a sound approach of the discipline Systems Engineering. Moreover, this attitude results in an excessive number of complicated situations involving the Systems Engineering staff members. In the past Dutch Education had enough motivated people to further develop Systems Engineering. As a result of the above staff-related problems enthusiasm has been reduced to zero. It is obvious that if the Dutch Government wishes to start-up new projects (such as

NIVO and PRINT) it should first of all clearly establish the manning aspect. The inefficient use of the means of production (educational tools) within Dutch Education has already clearly made its marks. As stated above, nobody has the status of being responsible or officially accountable for Systems Engineering. This is the main reason for the "proliferation" of the means of production. Any school may just "buy something".
Each and every staff member is sure to have his own views and this has resulted in an uncoordinated buying spree of computers and associated equipment. Education cannot expect any guidance from the industry and/or the dealers; they are exclusively interested in their sales figures. Obviously such a development has disastrous consequences for the financial position of a school and moreover for the discipline Systems Engineering. The above analysis shows that the present position of Systems Engineering is extremely complex. In order to start finding a possible solution to this problem it is of paramount importance to ensure that Systems Engineering becomes a formal discussion partner. As I have explained above, systems Engineering does indeed have interests although authorities and managements are only to eager to deny this. It would be in keeping with achieving that Systems Engineering becomes a formal discussion partner that we establish an interest group. The main object of such a group would be the organizational integration of Systems Engineering. It is clearly not the intention that this group would involve itself as regards substance. Obviously, contacts will have to be maintained at a managerial level with persons, groups and/or institutes in order to achieve an optimal coordination between hardware and educational targets. By means of this presentation I would like to communicate with interested parties or other parties concerned as regards how such an interest group should function at the various levels: local authorities, managements of schools, national authorities, teachers and other parties concerned. These discussions may lead to developing a new policy that will eliminate the present problems.

INTERACTIVE VIDEO IN SOFTWARE TRAINING-LEARNING WITH MULTIMEDIA

M.H. Jung, & R. Fricke, Seminar für Pädagogik, Braunschweig, Germany

The use of interactive learning- and video programs (computer based training) receives growing importance in education.
More and more companies are using interactive video for various training purposes.
The application of such programs promises a greater efficiency of schooling.
In order to know whether their use is effective, an evaluation of the increase of learners' abilities is necessary.
The purpose of this study is to show how interactive video can be used in software training, especially in word processing with Word 5.0.

The training concept

The concept was planned in such a way that the learners are able to appropriate the subject self-reliantly. The topics are presented in form of different auxiliary materials. Support is given by paper material and interactive video (i.e. multimedia) with the presentation of the topics and the learners' subsequent activity being fluently arranged. The use of auxiliary materials is voluntary. It is only necessary to apply the interactive video when help is really needed.
The different materials were created in order to increase the activity and the motivation of the learners. Furthermore, the concept contains tasks ranking from easy to difficult.

The conception with the different materials:

Written materials		**Software**
an exercises book		
		Word 5.0
a structural notebook		I
the original handbooks	learner	the interactive videoprogram
a paper which the learners receive first		

The software training begins with reading the paper which contains an instruction how to start the software program "Word". The learners start the program and then they activate the word-trainer (the interactive video program). They see a screen with four parts:

- firstly, two persons, an expertise and a novice, are shown in a film; this part is given to motivate the learners;
- secondly, the expertise can be seen demonstrating the handling with the different materials;
- thirdly, a test is included which examines the learners' preliminary knowledge;
- lastly, tasks are presented. When a learner activates this part he sees the same tasks as in his exercise book. With regard to his preliminary knowledge (have a look to the third part!) he may choose a task. In the beginning he obtains information about the task at hand. Afterwards he leaves the interactive video program in order to solve the task. In case of problems he may activate the "word-trainer".

The evaluation

The evaluation of the training program refers to:

- the direct increase of the learners' theoretical knowledge (multiple choice test);
- the direct increase of the learners' practical knowledge (the learning process is recorded by a computer tracking program); and
- the learners' acceptance of the training program.

On a poster the theoretical part, the detailed concept, the experimental plan, and the results will be presented.

GIRAFFE, A COMPUTER ASSISTED INSTRUCTION PROGRAMME

A.K. Boekhorst, & T. Groot, Department of Book, Library and Information Science, Faculty of Arts, University of Amsterdam, Amsterdam, The Netherlands

'Access to Knowledge' means access to information. To have access one needs to know what information sources are available and how the information wanted can be retrieved. Information technology and its products are penetrating with increasing speed and a compulsory way into government, business, science and education but also into day–to–day life of the citizen. There is a growing amount of information, information sources, information media which is retrievable over a greater geographic area in a decreasing period of time by a growing number of people via several communication channels. For most people ordinary skills are not sufficient any more and instruction on information retrieval and information management is needed, even for day–to–day life (PIN–codes).
If we restrict ourselves to 'Scientific Information', information used by scientists, we see that information is retrieved via several information channels: oral, written (printed, or digital). Formalised (written) information plays a specific function in scientific communication. The vast body of formalised information is supposed to represent the universal body of knowledge. To be able to retrieve relevant information, students have to know how scientific communication in general and for the relevant branch of knowledge is organised. What information sources are available on what information media and via what communication channels. We gather that more and more students are poorly informed over the possibilities to find relevant information and lack the skill of an adequate use of information sources. Bibliographies form the most general group of reference material.
Library instruction in general and instruction in information retrieval can vary from a global introduction to a series of instructions. Fjällbrant distinguishes several ways of instruction:

- Group instruction;
  lecture
  seminar/tutorial
  demonstration
  guided tour
- Group and individual instruction;
  film
  video tape
  tape/slide

651

audio tape/illustration
- Individual instruction;
book, printed guide
practical exercises
programmed instruction
self–instructional material
individual help

Results of these trainings vary according to the given instruction itself and to the amount of interest of the students. Many of them are not interested in the subject until, in a late stage of their study, they are confronted with the problem of informationretrieval. Computer–Assisted–Instruction as an aid for the instruction in library instruction and information retrieval has several advantages among which:

- it can be used at a moment chosen by the student him/herself;
- it is interactive: offers feedback, help, control;
- evaluation can be built in and printed (paper/file).

The instruction in the use of the 'Bibliographical Apparatus', has been part of the curriculum of the Department Book, Library and Information Science (UvA) for many years. In this programme attention is payed to the history of scientific communication and the organization of the bibliographic apparatus. Besides these lectures the students are trained in the use of the bibliographical apparatus by way of several tasks, which increase in degree of difficulty during the course. In the end students have to compile a small bibliography in which they have to describe how the bibliography came into existence.

In 1989 a two year project 'CAI & Humanities' was initiated. A collaboration of the Faculty of Arts and IBM Netherlands. In this project our proposal for the development of GIRAFFE (General Information Retrieval All Faculties for Bibliographic Education) was honoured. The main objective of the project was the development of a general retrieval programme for information on general bibliographies.

The programme has two aims:

- To inform the user on bibliographies, that might fulfil information needs;
- To give insight into the organization of the bibliographic apparatus.

The group aimed at consists of students, employees of universities and visitors of university libraries who want to consult bibliographies.

The programme is also developed specifically to give students a basic understanding of the organization of the bibliographic apparatus, and how to use it. Students are taken through logical steps in order to learn how to find relevant bibliographic references.

The programme has a menu–driven structure, which is developed by the COO–lab of the Faculty of Arts (UvA) and BIZON. The programme has a tree structure. It is possible to go directly from one description to another. The name or term is highlighted in the text. Each highlighted letter or term indicates a choice possibility. The choice can be activated by pushing the <enter> key.

The text of the programme was written in WordPerfect. There are now about 220 screens defined; 128 screens contain descriptions of reference material. All retrieval actions lead to a description of a bibliography or other reference material or to an instruction for further action via the help programme. The programmeme

does not contain the bibliographic information itself but the description may contain a reference to the place where it can be found in the library or information service where the programme is available. The programmeme is flexible and can easily be updated or filled with any bibliographic reference material and call numbers. The programme will be available for MS–DOS personal computers with or without harddisk on 3.5" and 5.25" diskettes.
The references that are included in the programmeme are selected from the 'Lijst van Algemene Bibliografieën' and the 'Overview of bibliographies and bibliographical aides'. Criterium for determining whether a title was included in this edition of GIRAFFE is the importance of the publication and the availability on large scale in theNetherlands. The descriptions of the reference material on the screens have a standard structure. An INDEX can be activated with the F3 function key. The list can be browsed with the arrow keys, or by typing the entry. If the entry is not in the index, the most proximate term will be Highlighted. Via the <ENTER> key the chosen page will be displayed. If there are more pages with the same index word, an overview is given with the titles of the pages.

Conclusion
GIRAFFE is developed primarily as an aide for lessons on the structure and use of the bibliographic apparatus. The computer is an efficient teaching device because of its ability to react immediately to the nees of the individual student for checking the successive steps through the bibliographic apparatus. Those checks can be made by going 'backwards' into the search strategy via the <ESC> key. At the same time, it is possible to get help and supplementary information over procedures and terms via the F1 function key. The programme can also function as an aide to any user for finding information over the bibliographic apparatus in any information service. The programme can function as a separate programme or can be integrated in a shell with information on more specialized subjects.
During the ECER 92 convention a demonstration of the programme will be given.

WELCOME IN 'GIRAFFE'
A RETRIEVAL PROGRAMME FOR 'GENERAL BIBLIOGRAPHIES'

You can use this programme to find bibliographic information on a subject and to verify or append data on publications. A help programme is available at all times.
Choose between:
- Search for publications on a certain subject:
  choose SEARCHING FOR LITERATURE ON A SUBJECT
- To verify or append data about a publication:
  choose VERIFY OR APPEND DATA
- Activate the HELP PROGRAMME
- For information on this programme and on searching in bibliographies choose F1

Each query can be converted into one of these choices.

Fig. 1
Opening Screen

UNDERSTANDING TECHNOLOGY IN SECONDARY SCHOOL: CAN THE TEACHER STAY AHEAD OF THE PUPIL?

Shane J. Blackman, University of Surrey, Guildford, United Kingdom.

Introduction to the Problem

What has been the consequence of technology, as the new major educational innovation? From our data, at the school level, senior management and classroom teachers argue that there has been an uncritical introduction of computers into schools. "The beneficial effects of computers were blindly taken for granted". For the classroom teacher questions about the relationship between technology, teaching and learning were left unanswered. The scale of the task to achieve teacher computer literacy has been vastly underestimated.

Method

The evaluation strategy on this national project has been to provide 'rolling feedback': direct to the eight regional projects and the TVEI Unit, in the form of both formative and summative reports. Data collection has been supplemented by interview case studies and questionnaires administered to both teachers and pupils on the effectiveness and implementation of IT in school. Overall the range of research tools has been as follows: in depth interviews, questionnaires, review meetings, school visits, team evaluation meetings and general ethnographic data collection.

Results

The aim of these projects is to develop teachers' skills and aptitude towards greater utilisation of information technology in promoting effective learning across the curriculum. We found limitations relating to the technological compatibility of hard and software which have been underplayed, especially in relation to the acquisition of 'state of the art' technology on an adhoc and ambiguous basis. Both the adoption of IT policies and the support given to staff development in IT by senior school management have begun to transform many of the earlier negative tendencies.

Conclusions

Teachers and pupils each respond differently to the challenge of working with technology, but both accept that the computer acts as powerful motivator. It is in the

very consumption of IT that we can become blind to its influence. As Raymond Williams (1974) pointed out, various technologies always exist in a myriad of forms, but they always require a social use. In this sense, it has been possible for us to examine how teachers are sensitive to pupil preferences towards particular learning styles. An individual may prefer one style over another when working alone, but if the person has a greater preference for working in a group she or he may go along with the group's way of working. Finally, there may be strong socio-cultural influences which overlay pupil perceptions of the efficiency of different learning styles. These perceptions are in turn influenced by the pupils' self image and gender identity of what they are good at, which shape their learning style preferences. It is not enough to just get technology into school or even plan its management. The real challenge that IT represents is in the opportunity to enhance learning as a social process on both sides of the pedagogic relationship. In the development of a framework to support teachers in their implementation of IT democratic principles, such as openness, reflection, and flexibilty are essential, but what becomes crucial is the empowerment of teachers to develop and to learn; empowered teachers most effectively empower students. Learning is a social process not a technological one (Blackman and Brown 1992).

Discussion

IT: Ordinary not Special

What evidence can the project put foward to suggest that the use of IT is now becoming more effective in school? The *primary* focus of attention will be on teachers' use of IT in terms of four strategies:

- Organisation and Management;
- Experiment and Reflection;
- Teaching and learning sstyles;
- Social and Psychological;

In general, the impact upon teaching staff involved with the project clearly has been considerable. Crucially, IT is no longer identified as being self evidently relevant.. One key way to achieve this has been to demystify the technology and also to clarify its purpose as an instrument to enhance learning. In the past the glorified self importance of IT only served to increase the negative legacy of the computer expert. Such celebrations left many teachers cold. The triumph that this project points towards is the recognition that computers are ordinary rather than special. Enhancing the friendly status of computers deminishes associated fears and inadequacies. It is upon this basis that our evidence points to increased use of IT by ordinary classroom teachers. As one teacher recognises: "I could see they were just another tool, teachers were dazzled by this IT thing . But it's just a machine that does particular tricks and the thing is to work out what the tricks are useful for in your particular subject."

Issues of Generation and Gender

The *secondary* focus is on the impact of IT upon the pupils. There are significant issues of generation at work here informally, teachers have told us that children, including their own, spend much time playing with computers. "They have their stuff under the TV, usually with a whole set of different wires and connections, with

some combined running from one machine to the next. I've got no idea what's going on, but the TV and video still seem to work". As teachers stress again and again: "kids are used to playing with computers", this leads to a sort of cultural monopoly by young people of the image of computers as belonging to them not to adults. There is a parallel here with modern pop music which is seen by young people as belonging to them; it is of their age and confirmed by a society that constantly informs us of "technological wiz kids". Children take this attitude to school and intimidate teachers, sometimes overtly, other times covertly. This not an imaginary problem, but something teachers encounter on a day to day basis. There is also a significant gendering of technology here; teachers seemed to feel (although they had not considered the issue in depth) that male pupils would possibly show off or attempt undermine teacher authority through interruptions that suggest the teacher lacks knowledge (Blackman1987). An example of this was explained as where boys try to dominate the lesson through side tracking. A male pupil will try two strategies, on the one hand to state to the teacher disconnected factual bits of computer knowledge, on the other to use the "hands on" approach and take the computer to the limit of his own knowlege. In the first example the teacher's confidence is undermined because the student appears to possess a huge stock of information. In the second example, the teacher's confidence is undermined because they may not possess the computer skill to return the screen to where it was taken by the pupil. In both cases through play and masculine bravado, pupils have suggested to the teacher that they possess more structured knowlege and understanding than that of the teacher.

References

Blackman, S. J. (1987). The Labour Market in School: new vocationalism and issues of socially ascribed discrimination., in Brown, P. and Ashton, D. [Eds] Education, Unemployment and Labour Markets, Lewis, Falmer Press, pp.27-56.

Blackman, S.J., & Brown, A. (1992). Framework to Support Schools in the Use of Information Technology to Enhance Learning, in Estes, N and Thomas, M (Eds) The Ninth International Conference on Education and Technology, Vol.3, Austin Texus, Morgan Printing, pp.1182-1184.

Williams, R. (1974). Television Technology and Cultural Form, London: Fontana.

656

THE IMPLEMENTATION OF TECHNOLOGY IN THE CALSSROOM: AN EXAMINATION OF TEACHER BELIEFS, PRACTICE, AND CONTEXT IN SHAPING MEANING AND ADAPTATION

Karen Lee Jost, Northern Illinois University, & Philip L. Doughty, Syracuse University, USA

Introduction

Within their own classrooms, teachers create learning environments that reflect both their assumptions about teaching and their preferred ways of working with students (Kerr, 1989). Contextual factors of the particular school and classroom also influence the teaching practices which are possible to implement. When an innovation is introduced, teachers will adapt it according to their beliefs, teaching styles, work environments, and their understanding of the meaning of the change. The many unique translations of an innovation complicates defining the innovation and measuring its effectiveness. Within the classroom, when implementing the use of technology as an instructional tool, it is also important to observe the different ways that the tool can be integrated into the curriculum and the teaching strategies which are effective in its use.

Olson (1981) discusses why teachers redefine an innovation which in turn raises questions as to its effectiveness. He states that not only is it important for teachers to understand the potential of the change, but it is also important that the developer understand the dilemmas which teachers face in their particular environment when implementing a change or innovation. Teacher dilemmas may arise due to beliefs, practices, roles, work satisfaction, and/or contextual factors. It is essential that designers and developers understand the significance of the changes they are asking teachers to make. Change efforts challenge well-established methods and their meaning. The meaning that innovators, designers, developers and teachers hold for both the innovation and school practice are generally not the same. "...it is in relation to existing goals, techniques, and social relations that teachers make sense of" innovations (Olson, 1980, p. 4).

With the emergence and use of new technologies in our society, questions arise concerning the need for change in our educational goals. We question the possible role of technology, both in shaping goals and in implementing reform. Mathematics reform in general and calculus reform specifically is looking toward the use of technology within the classroom. As with the implementation of any new curriculum we not only need to understand the goals of the change, we also must

657

view the way that individual teachers adapt the change within their own contexts and belief systems. We need to understand how the change is implemented at different sites and why the implementation is adapted in a particular way.
Past research on the use of technology in the classroom has generally focused on the "effectiveness" of using the technology as part of instruction. The recent history of the use of computers in education demonstrates that knowledge of the "effectiveness" of using a technology in the learning environment is not sufficient for successful implementation or for real change to occur. Adoption and acquisition are not enough. It is necessary to understand the changes that may be required, not only within the educational system but within individual classroom environments. We must therefore, view the change in terms of the teacher and whether the change necessitates a change in views or beliefs concerning learning and instruction or in teaching methods and strategies. We should also explore what is possible in real classrooms.
The context of this study is a pilot program which implemented a new curriculum in an introductory calculus course offered by Syracuse University's Project Advance in participating high schools for college credit. The new curriculum is centered around the integration of the use of programmable graphics calculators as an instructional tool. The new curriculum was not highly structured in terms of the daily integration and role of the calculator in instruction. The pilot teachers explored the uses and strategies for integrating the calculator. The teachers were thus an integral part of the process defining the new curriculum.

Purpose
The purpose of this study is to examine the process of implementation through the beliefs and practices of teachers in order to gain an understanding of the relationship between beliefs, practice and context in defining the innovation. The teachers' personal construct system and each individual's understanding of the meaning of the change must be considered in building an understanding of the perceptions of the innovation, of its perceived value, and its actual implementation (differing adaptations). This study attempts to uncover the meaning of what is being said or done, what Olson (1988) refers to as "thick" description, a term made famous by Clifford Geertz (1973).
This study incorporates some basic assumptions drawn from a combination of theory, past studies, observation, and experience. It is suggested that the strategies and methods used by teachers are influenced by a number of factors including their personal beliefs about learning, contextual factors, and experiences with their students. It is difficult for teachers to change their style of teaching, therefore, their teaching methods in the classroom may not readily reflect a change in their beliefs. Teachers with similar beliefs may construct and implement their lessons very differently. Also, teachers who appear to have very similar teaching styles may have different beliefs or intentions behind their actions.
The general questions shaping this study are questions which are important to describe the process of implementation, to aid the teachers in reflecting upon their practice, and to aid the innovator in understanding the adaptations of the innovation as it is implemented. These questions evolved from the literature on implementation (cf. Crandall & Eiseman, 1983; Akker, 1988; Olson, 1980), Kerr's

discussion of developing models for teaching with technology (cf. Kerr, 1989), and literature concerned with research on teacher's thinking (cf. Bussis et al., 1976). The following general questions shape the direction of this study:

- What are the teachers' constructs for teaching and learning?
- What contextual factors influenced instruction?
- How did the teachers actually use the calculator during instruction?
- What meaning did the teachers acquire for the change?
- What relationship(s) exist between teachers' constructs, contextual factors, actual use, and meaning?
- What are the outcomes of the change, implementation outcomes or instructional outcomes?

Answers to these questions define the innovation as implemented, describe the real change that has occurred, the degree of fit of the innovation, and its outcomes. The study also identifies factors which influence the use of the calculator as an instructional tool.

This study concerns learning environments which are undergoing change. In developing a conceptual model, it became apparent that there are many interdependent, inseparable variables in the study of how and what change actually occurs in a classroom. This study required what Salomon (1991) refers to as a systemic research approach. In order to gather teachers' perceptions, an understanding of their personal constructs and the intents behind their actions, in-depth interviews were conducted. These data were supplemented with fieldnotes from classroom observations, and observations and transcripts gathered at group meetings.

Methodology

The methodology employed in this study is phenomenologically-based analytic induction. It makes use of qualitative research techniques. The primary data were collected through in-depth interviews with each of the teachers in the Pilot Program, during May and June 1991. The interviews were audio taped, then transcribed. These data were supplemented with fieldnotes from classroom observations, memos from phone conversations with teachers, the analysis of documents (lesson plans, teacher journals, or other teacher-generated documents), interviews with the Project Advance Associate in charge of calculus, and the Mathematics Professor who initiated and designed the new curriculum. Additional data in the form of field notes were collected at the 1990 Summer Workshop for the pilot teachers, the 1990 Project Advance Fall Seminar for the pilot teachers, and the 1991 Spring Seminar for all calculus teachers. The Spring 1991 Seminar for the pilot teachers was audio-taped and transcribed.

The interview questions were designed as a framework for eliciting information about the background of the teacher, school and classroom context, the teachers' beliefs about teaching as well as their perspectives concerning the implementation of the new curriculum. The teachers' personal constructs concerning teaching are useful in understanding why they interpret and implement the innovation as they do.

The tapes were transcribed. The transcribed data and field notes were then coded. Coding categories developed from the data by identifying the concept that was illustrated by the data. The coded data were then sorted. Categories were further analyzed for commonalities or differences within categories. Categories which were often co-occurring were examined to determine if they were indeed separate categories or whether there exists a logical relationship between the categories and if so, the nature of the relationship.

The data was sorted retaining the label of where the data originated both to be able to go back to the original data in context, and to be able to analyze relationships and meaning for each teacher separately. This part of the analysis provides an understanding of each teacher's beliefs, environmental factors which may be shaping or constraining the teacher's actions, the meaning of the innovation which had developed for the teacher, and the definition of the innovation at the site (how it was actually implemented). This sorting method also allowed for a comparison between the different teachers. Thus, analysis was conducted within categories, across categories, within a teacher's data, and across teachers. It was also important to analyze the particular responses to some of the interview questions. Not only what was said, but what wasn't said was often important. In developing an understanding of the meaning of change for teachers it was also beneficial to observe the differences between how teachers actually used the calculator versus the uses that they were able to discuss.

Conclusions

The teachers did not make dramatic changes in their teaching styles. This is consistent with Olson's (1986) case studies of teachers integrating the computer as a teaching tool. The use of the calculator introduced changes in curriculum content and instructional method. The changes in instructional method included how some topics were introduced. These changes were generally implemented, however, without a change to the individual teacher's teaching style.

The teachers who have an interactive or inquiry style of teaching, used the calculator more. It appears that the use of technology as an instructional tool requires the ability to employ interactive or inquiry-oriented methodologies, which is consistent with Olson (1981). A traditional lecture style does not support the use of these methodologies. In addition, the capability of employing an instructional method appears to be dependent on holding compatible views concerning teaching and learning.

Meaning for the change developed with the teachers' use and experimentation and through their participation in seminars where they could discuss their experiences with each other. At the end of the school year, the meaning they had gained for the use of this technology in their classrooms ran along a continuum from perceiving it as a computational tool (mathematics tool) to that of an instructional tool. Those on the computational end stressed content-oriented goals and viewed learning as listening. The teachers who grew to view the strength of the calculator as an instructional tool have student-centered and discipline level goals for their students; interactive, inquiry-driven teaching styles; and constructivist views on learning. One of the teachers who saw the instructional tool strengths, also viewed the programming capability differently. He saw the benefits of the calculator as tutee.

The use of the calculator as instructional and learning tools has the potential of supporting more active learning environments. This potential not only interacts with a teacher's preferred style of teaching as stated by Olson (1984b), it also interacts with a teacher's goals, view of their role, and views concerning teaching and learning. Personality traits like patience and risk-taking may also influence initial use. In addition, the teachers who were reflective and critically analyzed their teaching were capable of perceiving the change that occurred in their teaching.
The use of programmable graphics calculators brought up the same issue that the use of computers in education has already raised. Should technology drive the curriculum or should technology be used as a tool for implementing the curriculum? I don't believe the answer lies in choosing sides. I think that we need to view the possibilities from a broad perspective which draws from both. The teachers' experimentation with the use of the calculator illustrates positive changes that took place that were due to the use of the technology and were not planned. The technology is already driving the curriculum. The question now should be in terms of how much the technology is allowed to shape the curriculum. With the experience and knowledge of how it can be used and what it makes possible, we need to step back and reexamine our goals for our students and what they need to learn in terms of the discipline. We can then come full circle by examining if we can help students to achieve those goals by using the technology as a tool. In this manner we can use technology as a medium for reform.
The implications from this study address the areas of teacher thinking, the implementation of educational change, the use of technology in the curriculum, mathematics education, and some basic issues of curriculum.

References

Akker, J.J. van den (1988). The teacher as learner in curriculum implementation. Journal of Curriculum Studies, 20(1), 47-55.

Bussis, A. M., Chittenden, E. A., & Amarel, M. (1976). Beyond surface curriculum: an interview study of teachers' understandings. Boulder, CO: Westview Press.

Crandall, D.P., & Eiseman, J.W. (1983). Coordinating assistance in school improvement efforts: issues to consider. Paper presented to the Annual Meeting of the American Educational Research Association, Montreal, Canada, April 11-15, 1983.

Geertz, C. (1973). The interpretation of cultures. New York: Basic Books.

Kerr, S. T. (1989). Technology, teachers, and the search for school reform. Educational Technology Research and Development, 37(4), 5-17.

Olson, J., & Eaton, S. (1986). Case studies of microcomputers in the classroom: Questions for curriculum and teacher education. Education and Technology Series,Ministry of Education, Ontario.

Olson, J.K. (1984). Microcomputers and the classroom order. Paper presented at the annual meeting of the American Research Association, New Orleans, LA, April 23-27. ERIC document ED277347.

Olson, J. K. (1980). Teacher constructs and curriculum change. Journal of Curriculum Studies, 12(1), 1-11.

Olson, J. K. (1981). Teacher influence in the classroom: a context for understanding curriculum translation. Instructional Science, 10(3), 259-275.

661

BEST COPY AVAILABLE

Olson, J.K. (1988). Through the looking glass: Towards thicker description of teaching. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA, April 5-9. ERIC document ED292799.

Salomon, G. (1991). The Qualitative Quantitative Debate. Educational Researcher.

'TECHNOLOGY-ENRICHED SCHOOL' PROJECTS

G. Carleer, & B.Collis, Department of Education, University of Twente, Enschede, The Netherlands

In many countries, 'technology-enriched school' projects have been established in order to study the effects of broadscale implementation of computer-related technology on all aspects of the school. What is the contribution of these projects to the more general understanding of the impact of technology on teaching and learning within complex social settings? Innovative approaches to research are necessary to adequately respond to the multi-level aspects of such projects. In addition, strategies to synthesize the experiences of these initiatives, and to enriched school projects and also schools in general are very much needed. In this symposium, researchers involved with such projects will reflect on the research directions proving to be most productive in their own projects. They will also highlight important observations emerging from the implementation activities in their schools. Following the presentations, the discussant will comment on the common themes and issues emerging from the research in the different settings. Participants will be asked to speculate on the transferability of their experiences to other schools and projects.

THE IMPACT OF INFORMATION TECHNOLOGY ON CHILDREN'S LEARNING

Margaret Cox, King's College London, London, Great Britain

For many years there have been various national and regional projects on the introduction and integration of information technology in schools, involving software development, hardware resourcing and teacher training. Although there have been numerous research projects to evaluate the impact of information technology on children's learning, there has also been concern in the research community that relatively little is still known about the effects of the use of information technology in schools. In order to evaluate such effects, in January 1989 a UK project was commissioned by the Education Ministry, and based at King's College London.

The project researchers studied a total of 86 technology enriched and non-enriched classes of pupils aged between 8 -10, 12 -14, and 14 - 16 years in English and Welsh schools, over a period of two years. This study involved both the assessment of achievement of particular subject based tests and in-depth case

studies of a subset of the 86 classes. The study focussed on pupils progress in English, geography, science and mathematics. Although the original intention was to work with experimental and control classes for the assessment, no two schools or classes remain stable over two years in their use of information technology. We therefore developed a rating scheme for the level of use of information technology in the class, which has been applied to every one of the 2400 pupils. This has enabled us to relate the change in their performance of the set tasks to the level of information technology use.
The results, still being analysed, indicate that only those pupils, which make regular and sustained use of information technology in a subject, show evidence of a sustained improvement in learning compared with non users. However, there is also evidence that specific use of a particular piece of software can also enhance children's understanding in that particular domain. The research results cover many assessment tasks, and classroom observations, and the methodology used could be adopted by other researchers wishing to study other age groups, subject groups, or different uses of information technology.

EVALUATION OF/IN TECHNOLOGY RICH PROJECTS - SUBSTANTIVE AND METHODOLOGICAL ISSUES

Berner Lindström, Gothenburg University, Mölndal, Sweden

In Sweden there has been a national program for the development of computers in education. While this program focused on didactical aspects in specific subjectmatter areas, some projects started from the idea of building technology rich environments as a necessary condition for this development and especially for broad scale implementations.
In Goteborg two projects were set up with this basic idea, including several partners; local schools, local school boards, a computer manufacturer (IBM and Apple), the university (teacher training and pedagogical research). The schools participating were chosen since they were "practice schools", i.e. schools where the teacher trainees practice, with interested teacher staff and school administration. The idea behind this was to build a bridge between the developmental work and the teacher training, where, among other things, the trainees could carry out their examination work in the schools, thus collecting data on the developmental process and of the utilization of computers for instructional purposes.
The IBM-project has so far been running for three years, and been the object of a larger evaluation study, where the project has been followed during its existence. The evaluation has been, what we call, a "pedagogical climate study", focused on a number of aspects: the computer as an instructional aid, the placement of equipment, motivation, concentration, learning effects, creativity, social interchange, a changing teacher role, and special education.
The Apple project has been running for a year and a half and has mostly been a platform for studies of learning in specific areas - arithmatics, system dynamics, physics, music, reading and writing. In all these studies we are looking at how the computer can contribute to a change in children's conceptual understanding. In

other studies we have been focusing on the pupils conception of the technology and on teachers conception of the technology and the pedagogical benefits.

RESEARCH IN TECHNOLOGY ENRICHED SCHOOLS: A CASE FOR COOPERATION BETWEEN TEACHERS AND RESEARCHERS

J. Bieshuizen, Department of Education, University of Leiden, Leiden, The Netherlands, & J. Moonen, University of Twente, Department of Education, Enschede,The Netherlands

Research into technology enriched education in three secondary schools in the Netherlands is conducted according to a strategy with four major characteristics.

- The project is an example of the *case study approach*. There are three experimental schools and no control schools. Two experimental schools are collaborating with one research team (Twente University). The third school is linked to a second research team (Leiden University and Utrecht University).
- *Close collaboration between teachers and researchers* is the most important characteristic of the research strategy. New applications are sometimes initiated by the research teams, but very often by the teachers themselves. In some departments, teams of teachers and researchers have been created to develop, implement and evaluate new applications. Consensus is sought for when experiments are planned and results are discussed.
- Effects are studied over a *wide spectrum*, including the classroom level (learning effects, changing roles of teachers and students), the curriculum level (changes in content, emphasis, sequence) and the school level (changes in school organisation, in attitudes of teachers) .
- The project has a relatively *long-term* character. Changes are studied over a period of four years, with many intermediate formative evaluations.

According to Dutch standards, the three experimental schools have been equipped with a considerable amount of hardware and teacher facilities. This raises the question whether the research findings can be generalized to other school settings, with less favourable circumstances. There are two arguments for an affirmative answer. First, the project produces a number of prototypical exemplers of successful courses in which computers are usefully integrated. These exemplers can be used as models for teachers in different situations to create their own technology enriched courses. The exemplers are carefully documented and distributed amongst secondary school teachers. Secondly, factors determining success and failure of computer applications are analysed in retrospect. These success or failure factors refer to hardware and software conditions, characteristics of teachers and subject-matter, and the institutional environment of the school. It is expected that the message which is learned by analyzing successes and failures, will be relevant for other schools.

DIDACTICS AND EDUCATIONAL COMPUTER SOFTWARE

A. Dirkzwager, The Netherlands

PREMISES FOR THE DEVELOPMENT OF HIGH QUALITY EDUCATIONAL SOFTWARE

E.J.R. Koper, Open University of the Netherlands, Heerlen, The Netherlands

Introduction

It is generally agreed upon that the effectiveness of educational software depends, among other things, on the systematic application of didactic theories in the software design methodology and on the degree of integration of the software with other instructional media in a course. One of the problems is that there is a lack of dedicated techniques for courseware development, integrating instructional design techniques and software engineering techniques.

We developed a design approach, called 'PROFIL' ('PROduction strategy For Interactive Learning environments'), in which we try to overcome this problem (Koper, 1989a; 1990; 1991).

The phases of PROFIL are:

- A preliminary investigation in which the learning objectives, student characteristics, didactic scenario and media are considered, defined and justified for a specific multimedia course under development.
- A definition phase in which a further didactic analyses and specification takes place for each of the different media applications to be developed in the course. For each application, including the educational software, the functional structure will be derived from a .didactic design, called the 'didactic scenario'.
- A script phase in which a full functional specification is made of the computer program to be developed. The script is specified in a special design language, calle 'INSCRIPT' (Koper, 1989a; 1990).
- A technical realisation phase. In this phase the program is technically designed, coded and tested, according to the functional specifications in the script.
- An implementation phase, in which the reproduction, marketing, training and distribution of the course, consisting of the different media applications, takes place.
- An exploitation phase. This includes maintenance, support and evaluation of the course.

PROFIL is the standard methodology for courseware development at the Dutch Open university. In the past years several dozens of educational software programs are realised with the help of PROFIL, ranging from tutorials in discrete

mathematics to simulation programs in biology and several different interactive video practicals, e.g. in the field of an 'assessment centre' approach for staff selection.
Because it is not possible to present PROFIL at length in this symposium, we will concentrate on the premises underlying PROFIL, which in our view are critical for the quality of courseware in general.

The premises of PROFIL

The first premise of PROFIL is that electronic media should be developed as an integrated part of multimedia courses and that the systematic selection, integration and justification of media applications is a major factor with respect to quality.
In our view a course consists of different media applications. A media application is described in terms of learning objectives, student characteristics (prior knowledge, learning needs and situational factors), a didactic scenario and a physical medium (e.g. teacher, text, computer and video). The didactic scenario is a imaginary, 'media neutral' design of a learning environment in which the student is thought to interact to attain the learning objectives. It is based on learning objectives, student characteristics and a didactic model, such as: (cognitive) apprenticeship learning (e.g. Brown, Collins & Duguid, 1989), problem based learning (Barrows & Tamblyn, 1980); open learning (e.g. Walberg, 1975); experiential learning (Kolb, 1984), etc. 'Media neutral' means that the functional objects in the hypothetical learning environment may be implemented in different available media.
Each media application is used for which it is suited best, taking into account the effectiveness, costs, feasibility and theoretical/ideological factors at the same time, but with different priorities (see Koper, 1989b). Clark & Sugrue (1990) discussed the problem that media as such do not have any psychological influence over learning, but that media may influence the economics (speed and costs) of learning (the so called 'weak media view'). The instructional design is a more critical factor. As mentioned earlier, a media application goes beyond the physical media itself. It also gives an account of the instructional design. Differences in terms of effectiveness, costs, etc. are therefore attributed to differences in the characteristics of different alternative media applications.
The second premise is related to the first: A development methodology should be compatible with different didactic models. The problem with some of the existing methodologies is that they are related to a certain didactic model and are not very well suited for other types of models. For instance, the development methodology presented by Bork (1984, 1986) is especially suited for highly interactive Socratic dialogues, but less suited for the development of practicals and open learning environments. Another example is the methodology presented by Burke (1982), which is solely directed towards the construction of programmed instruction.
The third premise is that it is possible to infer the overall design of the user-interface and the technical composition of the software, gradually, from the design of the instructional system. This premise is critical for the intended integration of instructional design techniques and software engineering techniques. Within PROFIL this is done in the definition phase. First a didactic scenario is specified. Then it is analysed in terms of the functional objects it contains, the interactions and the communication channels between the objects. These objects, interactions

and communication channels are then translated to implementation objects, interactions and communication channels, given by the available media.
The fourth premise is that the development of courseware is considered to be a kind of translation process in which an idea is translated into a script and a script into a computer program (Koper, 1990). In this context, it is important to use script languages and prototypes in the design phase to foster the communication processes in a multidisciplinary development team and to reduce the need of (cost-inefficient) feedback loops. One consequence of this view is that we do not expect that the use of authoring systems will lead to high quality software when they are used by teachers who implement their ideas directly into the machine.
Fifth, the development methodology should be compatible with different software analysis, design and programming techniques, specifically: structured programming and object-oriented approaches, but also e.g. the techniques used for the development of knowledge systems.
The last premise, related to the complexity of the work, is that the development of high quality software is teamwork, which involves different disciplines, such as educational technology, information technology, graphical design, subject matter speciality, audio/video production staff, etc.

Conclusion

For the production of high quality courseware we stated that it is necessary that the development methodology supports: working in multidisciplinary teams; the development of courseware as an integrated part of the media mix of a course; the systematic selection, integration and justification of media applications; the use of script languages; different didactic models and different software engineering techniques. These premises are underlying PROFIL, which we use successfully for the courseware development at the Open university. However, quality is one problem with the use of courseware. Another problem is to be found in the costs involved with the development and use of courseware. To some extent this is related to the efficient use of media in the media mix. When computers can do the job as effective as other media, including teachers, it sometimes is cheaper to develop and use courseware than to provide the education through e.g. teaching or text. This is especially true for institutions who work on a large scale.
However, a more important topic in this context is how we manage to re-use different elements of existing programs as much as possible for the development of new programs, without loss in design flexibility. Currently we are investigating this topic in terms of the concept of 'Flexible Educational Standard Objects'. This concept is expected to have an impact on PROFIL, which will be adapted according to the results.

References

Barrows, H.S., & Tamblyn, R.M. (1980). Problem-based learning: an approach to medical education. Springer, New York.

Koper, E.J.R (1989a). Inscript: een scripttaal voor het systematisch ontwerp van interactieve leersystemen. OTIC research rapport no. 4. Heerlen: Open univeristeit.

Koper, E.J.R. (1989b). Een keuzestrategie voor de inzet van (electronische) media in het hoger onderwijs. Tijdschrift voor Hoger Onderwijs,7 (3), 78-88.

Koper, E.J.R. (1990). Het ontwikkelen van Knowledge Acquisition Support Systems (KASS). OTIC research rapport no.16. Heerlen: Open universiteit.
Koper, E.J.R. (1991). Inscript: a courseware specification language. Computers & Education,16(2), 185-196.
Brown, J.S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. Educational Researcher. 18/1, 32-42.
Clark, R. E., & Sugrue, B.M. North American disputes about research on learning from media. Int. J. of Ed. Research,14(6), 507-520.
Walberg, H.J. (1975). Psychological theories of educational individualization. In H. Talmage (Ed.), Systems of individualized education.Berkeley: McCutchan.
Kolb, D.A. (1984). Experiential Learning.New Jersey: Prentice Hall.
Bork, A. (1984). Producing computer based learning material at the educational technology centre, J.of Computer based instruction, 11 (3),78-81.
Bork, A. (1986). Pedagogical development of computer-based learning material. In A. Bork & H. Weinstock, Designing computer-based learning materials. Berlin: Springer.
Burke, R.L. (1982). CAI sourcebook.New Jersey: Prentice Hall.

WHY COMPUTER SOFTWARE CAN BE EFFECTIVE

A. Schouten, Institute for Educational Measurement (Cito), Arnhem, The Netherlands

Introduction

In recent years a transition is taking place in thoughts about formal education. There is a move away from knowledge aspects towards the acquisition of higher cognitive skills like problem solving and information skills. The rationale behind this transition is that the skills required to take part in society are rapidly changing. In some professions new (production) technologies are introduced every few years, and also in private life new technologies seem more and more indispensable.

One realizes that knowledge and skills acquired in formal edcucation can no longer last a lifetime. People are expected to be able to acquire new material, new techniques in a short time. So students should (also) learn how to learn; i.e. how to formulate and solve problems, to gather and select information and to draw conclusions. This involves a substantial change in the teaching-learning situation. We think computers might help us there.

Roblyer, Castine and King (1988) reviewed 85 studies in their meta-analysis of research on the effectiveness of computers in education. They concluded that the effects of computer use in most studies are positive. Some of the results indicate that computers are especially effective for learning higher order skills. Roblyer e.a. conclude that with regard to type of application effect sizes are higher for simulations (science, e.s.= 0.64) and tutorials (language, e.s.= 0.41) than for drills. Many other studies are reviewed by Bialo & Sivin (1990). They report positive effects of computer use on student achievement, attitudes and social interaction. One of the studies that focusses on higher order skills, i.e. information-processing, is that by White (1987). He performed a two-treatment experiment involving a file-

management program. White mentions an effect size of 0.27. Another interesting study, focussing on effects on classroom organization rather than achievement, is the formative study around 'Earth lab' reported by Newman (1990).

Why computer use can be effective

To answer this question let us call in mind Bloom's "2-sigma" article (1984). On bases of several studies Bloom concludes that the most effective teaching-learning situation is one-to-one tutoring. In fact the results of the tutoring group was 2.0 standard deviations above the classroom control group. As a comparision: the effect size of mastery-learning was 1.0.

One reason why tutoring is succesful is the effect it has on the students' time on task. Time on task in the conventional control group was about 65%, in the mastery learning group 75% but during the tutorial sessions time on task was over 90%!

Another explanation for the large effect size can be found in the difference in interaction between the teacher and the learner. In a one-to-one situation there can be constant corrective feedback between the tutor and the learner.

Long ago this one-to-one tutoring was practiced in the guilds. The master guided and coached his apprentices individually, mostly during practice, until they themselves mastered the skills of the trade. Thus the student recieved a personal treatment suited to his needs.

Of course one-to-one tutoring in the sense of one-teacher-one-learner is not possible in our educational system: there are simply too many students. But with the help of computers we may be able to come close to that situation. There is some evidence from research that computers have the effect that students stay on task more persistently than in a classroom situation.

Furthermore computers, in combination with good software, can help the teacher approaching the ideal tutoring situation. As Newman (op. cit.) states: "Computers amplify the teachers capacity for a variety of kinds of instruction not by changing the teacher, but by changing instructional activity from, for example, one in which the teacher presents a book-based lesson to one in which the students work at a simulation in small groups". In that situation students can learn and practice at their own pace, while the teacher becomes their coach, who manages and monitors the learning process.

Why computers are still so little used

The research findings mentioned above show that computers can be effective tools in education, also for learning higher order cognitive skills. Then why are they still not very common in our classrooms?

We think it is because four important conditions for successful computer use in the classroom are not yet met: (1) availability of hardware, (2) availability of software, (3) teachers computer skills and (4) teachers didactical experience with computers.

- Availability of hardware

  At present secondary schools in the Netherlands have, depending on number of students, about 12 to 25 computers. 97% of these reside in a computer lab (Pelgrum, Ten Brummelhuis & Plomp, 1990). Often only a few teachers, the 'initiated', make use of these labs. The other teachers hardly use computers because of many reasons. Two of these reasons are: the labs are too often

occupied and there are too few computers outside the labs. So (many) more computers will be needed for useful integration of computer use in the curriculum to become possible.

- Availability of software
  If computers are to be used widely high quality educational software must be available. Software that suits the needs of education. This means that enough software must be available relevant to the acquisition of higher order cognitive skills.
  Unfortunately there is not yet much of this software. The 1988 OTA report (Power On! new tools for learning) showed that at that time over 90% of the computer software in the US had to be considered of medium or even low quality. Many years after the first introduction of computers in schools much educational software still showed programmatic flaws. The didactic quality seemed hardly any better. Furthermore, of the remaining computer software many programs were typically suited for practice and drill only. As Maddux (1991) concludes: application of this kind of software on a large scale might *increase* rather then *decrease* the emphasis on knowledge aspects in education. And that is not what we think would be desirable.
- Computer training for teachers
  Many teachers that teach nowadays have had little experience of computer use during their education or teacher training. Special training will be needed to make them feel confident with computers, before they can be expected to use computers in the classroom.
- Teachers Training
  Using educational software of the tutorial kind, or simulation programs, changes the role of the teacher. Teachers should be trained to be flexible coaches of the learning process rather than the traditional 'instructor'. Teachers must be capable of switching between roles: manager, tutor, peer, poser of questions. Often teachers become an interested learner with the students. This, of course, means a significant change from how they have learned to teach. On the other hand many teachers who work with computers report to be more aware of and feel more involved in the learning process of the students (OTA report, p.89 a.o.).

Conclusions

As the demands of society change education should focus on the development of higher cognitive skills rather than knowledge aspects.
From Blooms 2-sigma article we know that the best learning situation is one-to-one tutoring. Of course this system is too expensive for formal education. But we do have computers, providing us with a tool that -within a few years- may be widely available. There are indications from research that computers in combination with relevant software may be effectively used to enhance the learning of higher cognitive skills. But to achieve that goal many things in the traditional classroom routine may need changing.
Right now, in 1992, the use of computers is not yet integrated in the curriculum. Four conditons are mentioned that should be fulfilled for computer use to be successful. We think the most vital of these concern the availability of relevant high quality software and the training of the teacher. Like the master in the old days the teacher should become manager of the learning process and a coach for the

student. At present many teachers are not yet well equipped for that role. So a lot of work is to be done. But if we are right it may be well worth while: the pay off may be (almost) as high as with one-to-one tutoring, i.e. 2-sigma. And that should at least be worth trying.

References

Bialo, E. & Sivin J. (1990). Report on the effectiveness of microcomputers in schools. Washington D.C.: Software Publishers Association.

Bloom, B.S. (1984). The 2-sigma problem: the search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher, 13, 6, pp.4-16.

Maddux, C.D. (1991). Integration versus computer labs: an either/or proposition? Educational Technology, 1991, 10, pp. 36-40.

Newman, D. (1990). Opportunities for research on the organizational impact of school computers. Educational Researcher, 19, 3, pp. 8-13.

OTA (1988). Power On! New tools for teaching and learning. U.S.Congress, Office of technology assessment. Washington, D.C.: Government printing office. Catalog card number 88-600551.

Pelgrum, W.J., Ten Brummelhuis, A. & Tj. Plomp (1990). The availability of instructional computer materials by teachers. Results from the Dutch participation in the IEA COMPED study. Paper presented at the 1990 AERA meeting in Boston.

Roblyer, M.D., Castine, W.H. & King, F.J. (1988). Assessing the impact of computer-based instruction. A review of recent research. New York/London: The Haworth Press.

White, C.S. (1987). Developing information processing skills through structured activities with a computerized file-management program. J. of Educational Computeing Research, 3, 3, pp. 355-375.

FORMALIZING HUMAN EXPERT KNOWLEDGE ABOUT FRACTIONS LEARNING: THE CONSTRUCTION OF A DIDACTICS EXPERT SYSTEM

Johan Zuidema, & Eveline Helsper, Freudenthal Institut, Utrecht, The Netherlands

One of the crucial parts of an Intelligent Tutoring System (ITS) is a Didactics Expert System (XS). Construction of a Didactics XS has to be an interdisciplinary enterprise, including at least contributions of domain-specific didactics, instructional science and computer science. In this paper we describe the way we develop a Didactics XS for fractions instruction. Didactical research, carried out without the objective of computerized applications, gives direction to the process of designing the Didactics XS. Computers, however, require complete, formal specifications of the exercises, i.e. specifications that cannot be formulated on the base of only earlier performed research concerning fractions instruction.At mapping out the didactics expertise needed for an ITS-fractions, we conduct two complementary procedures. On the one hand, we retrace the human expert

knowledge that is implicitly present in four lesson-books, crystallized in the kind and order of fractions exercises. The instructional knowledge of its different authors is systematically investigated by statistical research of all kind of mathematical characteristics of occurring fractions -- more or less a bottom-up approach. On the other hand, we systematically interview four experienced human experts -- the top-down counterpart. The first source of didactics knowledge leads immediately to formal, for a computer processable, didactics information. The information as elicited by the human experts, however, is often not directly useful, because it is partly incomplete or even mutually conflicting. Nevertheless, we consider this direct source of expertise as the most significant guide-line for the Didactics XS, because of its explicit character.A computer cannot itself handle conflicting information. But, we do not want to fully disambiguate the compiled didactics knowledge. We want to discuss some ways of escaping from this impasse, using recent techniques of Artificial Intelligence. The result is a system that generates 'best' or 'near best' didactical Plans, doing significantly better than a random generator.

COMPUTER-BASED TUTOR USING CONCEPTUAL STRUCTURES

Alla Terekhina, Institute for System Studies, Moscow, Russia

Acquisition of knowledge and skills is accompanied by forming cognitive structures in one's mind, which reflect the understanding. Cognitive structures are a mirror of knowledge organization in the individual's mind.
A skilled person keeps knowledge in his mind as a system while an unskilled one keeps it as a set of non-correlated items. Considering the state of conceptual structures one may distinguish stable knowledge from vague one, hence an expert from unskilled persons. When acquiring knowledge and skills a human being goes through three stages. First, he views the domain as a set of non-correlated items. At the next stage, he considers the domain in terms of these items but, this time, integrated into small and, later, bigger groups. At the highest level of skill, he develops a language of system generation (organizing dimensions) which gives him an opportunity to organize his knowledge into a system. Both organizing dimensions and cognitive structure are implicit. They reflect the memory design and cannot be specified in a direct way. Processing expert's estimates with MultiDimensional Scaling the computer-based analyzer may elicit some implicit structures of expert's knowledge. This provides an empirical way for revealing organizing principles of knowledge structures.
So in terms of conceptual structures acquiring knowledge and experience is a way from a set of non-correlated items towards a system. A computer-based tutor was developed to assist a subject to systematize his knowledge in the memory. It uses expert's conceptual structures for tutorial purposes. It explains and specifies conceptual structures of skilled experts both graphically and verbally. All the information is elicited from experts.

For example, let us consider two conceptual structures, which are used for tutorial purposes. The first topic is connected with MDS area. The methods are coded by the author's names.
All the methods integrated into a few groups according to the main MDS models. The first dimension is associated with the data sort (it is omitted in the figure). Coombs' and Roskam' algorithms are designed for preference analysis. All the rest is proximity analysis. The second dimension divides the methods according to the number of subjects involved into the experiment. Coombs', Roskam', Carroll' and Horan' algorithms deal with group data. All the other process single data. The third dimension is associated with the type of mapping (metric, semimetric and entirely nonmetric). Torgerson is entirely metric, Coombs is entirely nonmetric. Kruskal, Guttman and Johnson are in the middle of the scale. More detailed structure corresponds to a some narrow area.

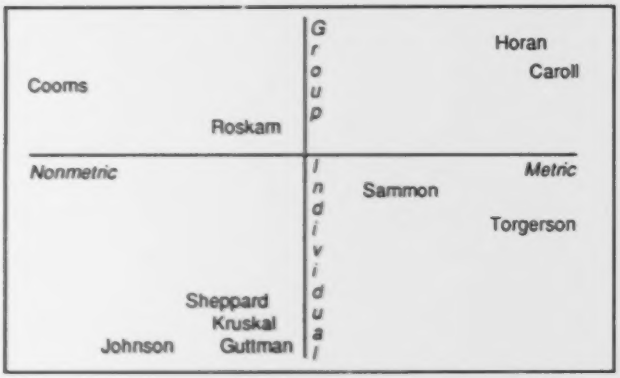


Fig 1

Another topic is English verbs such as must, to have to, to be to, ought to, should, would, shall, will, to be obliged to. In the experiments a few English speaking people were involved.
Figure 2 shows a joint semantic space for these persons. The horizontal dimension is associated with the intensity of modality. On the right, the action will be done by all means. On the left, the action may be undone. The vertical dimension is associated with the source of modality (wether it is inner or outer). At the bottom, one is forced to make the action by another person or by circumstances outside his control. At the top, the action is initiated by one's own will.

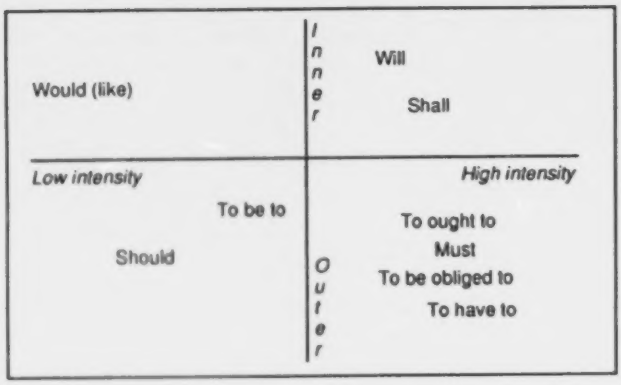


Fig 2

First the subject is tested and his cognitive structure is examined. The computer-based analyzer explores his understanding of the topic's structure. To test the subject it is necessary to specify the level of his conceptual structure, which reflects his understanding of the topic. Such analyses may be conducted for the constructed for the field as a whole and parts of it. If the computer-base tutor finds out that the subject's understanding of the topic is inadequate (the subject does not take into consideration the organizing dimensions, which are principal for knowledge systematization), the subject is provided with the missing knowledge about the topic's structure. Then the analysis of individual concepts (elements) is made. If the tutor finds out that the concept does not take a proper place in the system, its specification in terms of organizing dimensions is given to the learner. The explanations are delivered in terms of cognitive structures, which reflects thinking, that is why they are perceived rapidly and rather easily.
After the tutorial session one can use the organizing dimensions developed by experts in order to systematize the knowledge in one's mind. The computer-based tutor assists to organize a stable knowledge system in one's mind and thereby improves the understanding of the concepts under study. In this mode it is suited for several educational purposes.

SPECIFICS OF HUMAN THINKING UNDER COMPUTER CONTROL

L. Babanin, Department of Psychology, Moscow State University, Russia

The difference between computerized and traditional didactics is not only in the fact that computers teach using strong algorithms but also in the fact that human thinking in dialogue with other human is quite peculiar compared to the process of

thinking in dialogue with computer. In our experiments the process of knowledge reproduction in dialogue in the form of giving various definitions. to well-known concepts was investigated. We used two forms of dialogue: with computer and with other human (experimenter). The dialogue scripts were the same in, both variants, their aim was to help the subject to reproduce the knowledge. At the first stage of the dialogue script the subject reproduced the knowledge without any actions, when he refused the second stage was beginning. At this stage the subject encouraged to perform mental actions using the concept: comparison, abstraction finding of similar concepts and antonyms etc. As an indicator of effectiveness was chosen the number of concept attributes which subject produced when making definitions.

At the first stage of the experiment subjects produced more attributes in the dialogue with experimenter than in the dialogue with computer. But the following mental actions under computer control were much more effective than under the control of experimenter. As a result in the dialogue with computer subjects produced no less attributes than in the dialogue with experimenter. But the results show that knowledge reproduction did not submit the action directly. The orders given by computer did not define the answers of the subjects. It activated human thinking process concerning the answer. The concrete answer might have nothing to do with the order, it might be connected with general knowledge about the concept.

Thus the psychological specifics may be constructive when it is used in computer systems in the form of the didactic principle: do not evaluate subject's knowledge only by the first reply, in the case of bad answers there are recommendations to encourage the subject to do some mental actions with the knowledge and repeat the answer.

INTERFACE DESIGN FOR EDUCATIONAL SOFTWARE

Jan van der Linden, & Egbert M.H. Assink, Department of Psychology, University of Utrecht, Utrecht, The Netherlands

Introduction

The research project 'Feedback Processes in Computer Managed Spelling Instruction' is aimed at designing and testing courseware for teaching the complex rules of verb orthography in Dutch elementary and secondary schools. The courseware has been designed to enable the student to master the subject matter without the help of a human teacher. The computer program performs the main tutoring functions, such as presentation of the learning content, supervising the training sessions and evaluating the advancement in the learning process. Student controlled options for progression tempo and instruction/feedback level have been provided. The progression through the software is constrained by the program's evaluation of student results.

One major goal of the research project was to investigate the effectiveness of the courseware, taking into account individual differences in spelling and negative fear of failure. The effect of the courseware on spelling performance was measured by means of a pretest-posttest design. In a field experiment the

courseware produced positive results (Van der Linden & Assink, 1990a, 1990b). Students with high spelling ability outperformed students with low spelling ability, but both reached a significant learning gain. Students with a low fear of failure did not outperform students with a high fear of failure. Overall the students with low spelling ability and high fear of failure used significant more learning time compared to their counterparts.
Since the target user group was inexperienced with computers, special attention had to be given to the interface design. Although instructional design theories at a general level do provide some suggestions, there were no clear prescriptions about how to design the computer interface six years ago. Therefore the design team was obliged to develop their own ideas in this respect. In this paper we will evaluate some decisions with regard to the interface design. In evaluating these decisions, recent publications on research in interface design were used. The paper presents recently suggested interface design criteria and discusses the question if these criteria enable us to further improve the interface of computer instruction.
The criteria are grouped by four processes that are essential to the development and use of courseware. The first process concerns the macro level and concerns choosing the design method and selection of the developent team. The other processes represent the micro level: the course of the student-computer interaction process, optimizing information processing from the monitor and the enhancement of educational effectiveness.

Design Strategy of the Project
Developing courseware is a complex activity demanding expertise at a variety of design specialisms. At present the research literature is conclusive on this point. Courseware design and development is decisively not an undertaking which can successfully be accomplished by individual effort. Expert knowledge with regard to the task domain, as well as expertise in informatics should be present in the design team. Moreover, recent experiences collected in related research projects suggest that design expertise in the domain of traditional informatics should be supplemented by expertise on human factors design, in order to develop highly interactive software.
A major feature of the design method is user participation. In the early days a 'designer-centred' method was a common approach. The designer composed, tested and distributed the software entirely by himself, without any participation of representatives of the target user group. Recent literature claims that 'user centred system design' is the ultimate approach. A early and continuous participation of the user in the design process is strongly recommended, but in practice this advice is not often followed. Nevertheless, at this moment opinions converge on the point that designing software cannot be an enterprise that can be successfully accomplished in complete isolation from the target user population. Designers must somehow collect empirical information on functionality and usability demands.

Design of the Student-Computer Interaction Process
Courseware is often developed with the aim of teaching a skill or task that must be performed in the 'real world', that is in a non-computer setting. This is also the case

in our project. A major problem in this type of instructional context is that the student is in fact confronted with a double task. First of all, he must learn to handle the computer, interacting with the machine technically and, secondly, he must assimilate the learning content. This double task introduces a serious mental load problem on the user's side. It is evident that the learning content is the most important element for instruction. So, the course of the student-computer interaction should be as smooth as possible. The ultimate ideal is that the user experiences a feeling of complete control in a working session. This will reduce the detrimental effect of cognitive load

A smooth man-machine interaction can be accomplished by emphasizing the symmetry aspect of the dialogue. A user proceeds and reverts in the program, he navigates left and right, all possible operations should be revertible. Moreover, the user operations should be basically simple and arranged in an orderly format. The interaction must allow for flexibility. The subject's feeling of control is annihilated when the planned interaction dialogue only permits progression in a preplanned, rigid way. To illustrate, menu based interfaces should also be controllable by keyboard codes. Moreover, the courseware should present an appropriate response to the user input.

Information Processing from the Screen

Courseware consists partly of instructional materials that has to be read from the screen. Different types of information located on the screen are needed to communicate with the computer and to understand the learning content. This information must be presented in a consistent screen layout where the functional areas are clearly distinguishable. Consistency is also needed on the points of colour use and various text aspects. Processing text from the screen is often compared with processing text from printed materials. Until not long ago, reading proved to be much slower from a CRT screen. However, when nowadays text is presented on a high resolution monitor, in upper and lower case, double spaced, with proportonal spacing within words, with no justification on the right margin, and indenting each paragraph, the processing advantage of printed materials disappears.

Educational Effectiveness

Many aspects of teaching determine educational effectiveness. In a computer setting important aspects are instructional design strategy, feedback and help facilities. Preferably the student's overall performance measures, collected in a non-computer setting, should be considered as the ultimate test of courseware effectiveness.

Though it is impossible to prescribe an overall instructional design strategy for computer settings, the main objective is that the learner should exploit the full potency human-computer interaction offers, as compared to other instructional media. As a matter of fact, the manner computerized practice sessions are designed should express the unique characteristics of the medium.

Feedback research suggests the value of both direct and delayed responses in a practice session. Though feedback should inform the student about the cause of errors, it should never grow too lengthy. Each elaboration must contain new information. Informational feedback referring to a rule or definition has proven to

be successful. If possible feedback should be adapted to the learner style and task characteristics.

Proving help is another main tutoring function. Research indicates that good help facilities are associated with high quality software. Users must at any time be able to ask several kinds of questions. They may be procedural (how?), declarative (what?), causative (why?) and verificational (is this idea correct?) in character. But then some serious problems arise. Ideally the help system should assist the user to formulate the problem. If the student can formulate the problem in his own words it has to be 'translated' into application technology. Next, the student has to know if the present help system provides help on his problem, and if so, how to get it in the shortest way. Designing a help module satisfying these demands will take some effort and time.

References

Linden, J. van der, & E.M.H. Assink (1990a). Designing Computer Aided Instruction. In: Performance Improvement Quarterly, 3(2), 65-76.

Linden, J. van der, & E.M.H. Assink (1990b). Designing Instructional Man-Computer Interaction. In: L. de Leeuw, J.M. Pieters & P.R.J. Simons (Eds.), Computers in Learning and Instructional Research. Lisse: Swets & Zeitlinger / Dutch National Foundation for Educational Research.

The Computer As A Tool In Teaching Mathematics

Thomas Lingefjärd, Department of Didactics and Mathematics, Teachers College, Gothenburg University, Sweden

During the last ten years the use of computers in the subject mathematics has changed from pure programming and studies of isolated episodes to use of mathematical tools and specially designed software aimed to train a special part of mathematics (e. g. multiplication). We are now trying to change the studying of mathematics in our school system from routine work to constructing a conceptual structure and to problem solving. The routine work (e g manipulation) should be taken care of by computers and calculators.

Some main questions are:

- In what way can the use of a computer as a tool help us deepen the conceptual understanding and increase the ability to solve problems?
- What goals should be stipulated as reasonable, considering the power of a modern computer?

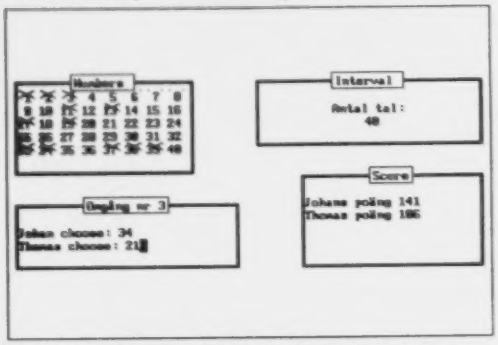
To investigate these questions, I have done three different case studies, in elementary school, in adult school (high school level) and in teachers college. The case studies all try to answer the main question: can we teach mathematics in another way with the help of a computer? Can we reach new topics and areas in mathematics if we leave the routine work to the machine?

Case study 1

Starting with elementary school, is it possible to introduce children at age eight to nine into number theory?

The case study has been accomplished with the help of the software called "Lek med Tal" (Play with Numbers), recording the discussion between a young boy (9 yr.) and the teacher. The game is an old mathematical game and is sometimes called Factors: you pick a number in a given interval (e. g. [1,..., 40]). You receive points of the same magnitude as the number, the person you are playing with obtains the sum of the true divisors of the number. Each number can only be used once. If, for instance, you pick 12 and receive the score 12, the other player scores 1 + 2 + 3 + 4 + 6 = 16. Obviously 12 is a bad choice, since it is a rich number. The best number to pick is evidently a prime number and after one player has done so, the number 1 is gone and no more prime number can be selected.

The result of the study was that even a 9 year old, could easily understand the strategy and so to speak learn number theory the natural way.



There is of course no need for a computer here (this could be done by hand), but the computer could deal with the administration, count the scores and so on. If you play against the computer instead of another player, you find that the computer is cheating! When no number could be taken any more, all leftover scores go to the computer.

Children normally find this very challenging (in a positive way) and very soon detect strategies how to beat the computer. College students examining this software, come to the conclusion that the software is unfair! Undergraduate teachers don't find the situation challenging at all, only frustrating. Their opinions and views will be discussed during the seminar.

Case study 2

From the national assessment concerning mathematics in part of our high school curriculum, we have gain information about the difficulty of different problems:

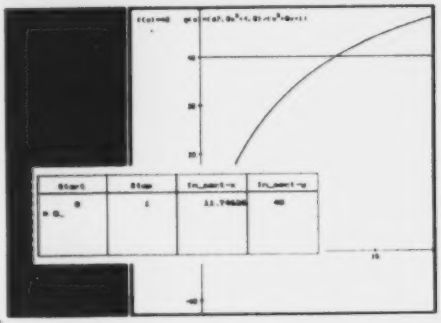
- Solve the equation $f'(x) = 0$, if $f(x) = x^3 - 3x^2 - 24x - 8$, has the solving frequency 0.60.
- The temperature fall y °C, in a water reservoir describes by the function $y = f(t)$. The letter t stands for the time in hours, counted from 8.00 o'clock, when the reservoir was filled.

One knows that $f'(3) = -4.2$. Describe in words what the derivative value $f'(3)$ tells you about the changing of temperature in the reservoir. Solving frequency 0.10.

Obviously the training of manipulating routine skills don't provide the students with ability to interpret the meaning of the derivative value in an applied math problem. We therefore investigated if unsure students can solve advanced math problems with the help of a computer? From a case study in an evening adult school we

know that students who seem to lack enough manipulating skills, nevertheless are able to solve problems in mathematics, equipped with sufficient tools.

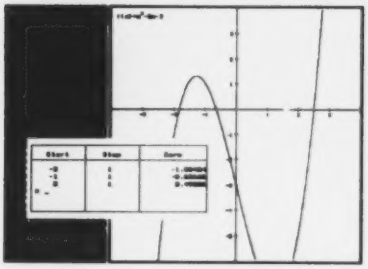
Michaelas father, who is a teacher of mathematics, says that Michaelas weight y kilo could be described by the equation:
$y = (67.5x^2 + 4.5)/(x^2+8x+1)$
where x is age in years.
At what age was Michaelas weight 40 kilo? The student:
I use the Work-shop and transform the prob-lem into two functions, sketch the graphs and seek for the intersection..

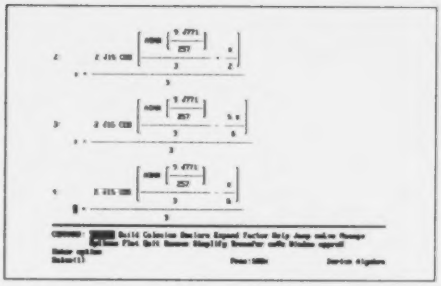


It may be discussed if this is to be considered good in mathematics, which is done in my paper.

At Gothenburg University we educate teachers in mathematics for elementary, primary, secondary school and teachers for our high schools (gymnasiums). They study among other subjects pure mathematics, methodology and do workshops of different kinds with the assistance of computers.

Case study 3
What kind of problem solving strategy could we expect from our students when they are given access to computers and other facilities (calculators). How will they solve the cubic equation, for instance? A recently finished study concentrated on problem solving focus upon attitudes and solving strategies. One introducing problem was: Solve the equation $x^3 - 5x - 3 = 0$.
I also asked the students to reflect about the meaning of *solve* the equation? Is there any hidden meaning in the question, that forces us to use a particular method?
A method to solve the cubic equation was first printed in Ars Magna, written by Girolamo Cardano (1545). Isaac Newton presented a numerical approach and method in 1669. Since then many numerical methods of solving equations have been published. The technical evolution of our time calls for some new methods and some old methods in a new shape. Most of the students (85%) considered the help of a computer to be an inaccurate method. From their earlier studies they have the opinion that an "exact" method is what the teacher is asking for. The two figures presents different kinds of solutions, one based upon the connection between roots and intersections (zeros), the other one based upon the Cardano algorithm.





The students were asked: *What kind of solution do you think you understand, or like best? What conceptual understanding do you need to interpret the two different kinds of solution?* In my paper, I will try to illuminate the difference in opinions among the students and discuss if it's proper to use a computer in the solving process. I also will show that we can reach further in problem solving, when leaving the routine work to computers.

Preliminary conclusions

"The older we get, the harder to change." This seems to be true also in the view upon mathematics as a subject. College students have doubts if it's proper to use a computer in the problem solving process. Younger students seem to look at it the other way: why should we not use the machine if it can make problem solving easier for us? Probably this is an effect of the educational system, ambitious teachers who implement rigor in mathematics too early and too well.

References

Greger, K., & Lingefjärd, T. (1989). Datorn i Matematikundervisningen I. Nämnaren, 16(3), (34-39).

Greger, K., & Lingefjärd, T. (1990). Ekvationslösning I. Datorn I Utbildningen, 4(2), (38-41).

Grevholm, B., Hedrén, R., & Lingefjärd, T. (1991). Datorn i den grundläggande matematikundervisningen. Tal och Räkning 2. (141-167). Studentlitteratur.

Lindberg, L., & Lingefjärd T., Visualization and rigor of proofs, unpublished manuscript, Gothenburg University 1992.

Lingefjärd, T. (1990). The Role of Computers in Teaching -especially Mathematics, Paper presented at the 2nd International Olympiad in Informatics, Minsk.

Report from department of Didactics and Mathematics, 1992:01, (Eds): (Emanuelsson, G., Lingefjärd, T., & Johansson, B.) Gothenburg University 1992.

HOW TO EVALUATE MULTIMEDIA SIMULATIONS: LEARNING FROM THE PAST

Rauf Yildiz, & Madeleine J. Atkins, School of Education, University of Newcastle upon Tyne, United Kingdom

Abstract

The main purposes of this paper are:

- to identify the main criticisms of media research;
- to discuss the state of current IV simulation research in the light of the criticisms made;
- and to identify the most important points which need to be considered by media researchers in order to avoid making similar mistakes in the future.

Introduction

Multimedia technologies hold out a great potential to educators. Interactive video (IV) for example comprises quality still & moving pictures, perfect digital sound, fast and accurate data processing power, and effective image and graphics display. Its potential for creating simulated environments seems especially promising. IV simulations are, also theoretically, capable of meeting the requirements recommended by learning theories for effective learning processes to occur. However, in practice, IV and multimedia simulations may not live up to expectations and empirical evaluation studies have themselves been repeatedly criticised (Clark, 1983, 1984, 1985; Salomon & Clark, 1977; Clark & Salomon, 1986).

The main criticism of media evaluation studies

Educators have criticised, since the 1950's, the way that media evaluation studies have been carried out.

1950's and early 1970's

In the 1950's and 1960's when a comparative evaluation format was widely used most studies reported no significant differences between experimental groups using new technologies and control groups.

criticisms of these evaluation studies were manifold. They included:

- producing 'uninterpretable' findings,
- producing 'meaningless' data,
- not producing enough information to determine the unique features of various media and whether they were effective in teaching different skills in various situations.

Other limitations of the comparative approach were identified as:

- inadequate treatment definition,
- incomplete specification of treatment dimensions,
- failure to measure treatment implementation, and
- insufficient relevance to practical decision-making (Reeves 1986).

From the 1970's to the Present

Research has been focused on the effectiveness of computer and multimedia applications. In the mid 1970's a new approach to media evaluation emerged. It

was called Aptitude-treatment interaction (ATI) and its main assumption was that learning involves an interaction of task, learner, and media characteristics. The phrase (ATI) was coined by Cronbach and Snow (1977). Have researchers followed the ATI approach and abandoned the comparative format? Although researchers who designed CBI studies were urged to avoid gross comparisons (Clark, 1985), the literature reveals that evaluators continued to conduct large numbers of studies using the comparative format, hardly focused on ATI, and progressed very little from the 1950's and 1960's. Indeed, one of the major sources of support for the use of computers for instruction has been the results of comparative research studies (Clark, 1985) which claimed increased efficiency for the new technology. It could therefore be argued that the same comparative approach is still being applied for media research and that the findings probably deserve the same criticisms.

Researchers who did use an ATI approach to media studies again reported substantial numbers of non-significant findings (e.g. Cohen et al., 1981). According to Moore et al. (1986), this was due to problems involving design, taxonomies, analyses, and a variety of 'real world' dilemmas.

The experimental design behind ATI and the heavy reliance on it have, in turn, been criticised (Bates, 1981; Reeves, 1986). Reeves (1986) concluded that research studies employing experimental and quasi-experimental designs to compare instructional technologies have produced few useful results and thus they should be abandoned. Salomon & Clark (1977) pointed to the lack of research carried out in the real life educational world, conducted with *real* materials. While the laboratory-controlled experimental method allows a better control over variables which leads to a better internal validity, and allows better conceptualization and understanding, it lacks representativeness, and hence, has only remote relevance to educational practice (Salomon & Clark, 1977). On the other hand, studies carried out in the real world of education, dealing with complex variables, are most often highly specific and do not warrant generalization. They also have a poor internal validity because of the complexity of the phenomena they deal with. Salomon & Clark (1977) conclude that internal validity may be sacrificed if the study is a summative evaluation, which involves 'judgement of worthwhileness'. The sacrifice is worth making because it allows media being studied to be used to its fullest advantage or to be exploited fully.

The evidence available on instructional effectiveness of IV simulations

The main findings on instructional effectiveness of IV simulations, based on an extensive literature review, can be identified as follows:

- most of the studies report positive attitude towards IV simulations;
- simulations are among the most educationally successful forms of interaction in an IV package;
- when IV is integrated in secondary school science lessons effectively it can improve standards of laboratory work and save time in comparison with setting up normal experiments;
- tutorial IV simulations in chemistry can significantly increase student performance on laboratory reports and test scores in comparison to those who do not;

- there is no significant difference between the cognitive effect of IV laboratory simulations and standard laboratory instruction in teaching college level pure science;
- IV's visual images can help college students to understand some science principles better;
- IV simulations can enable students to visit a real place and move around in time and space to investigate, sample, analyze, and test possible ideas within the environment.

Conclusions and recommendations for future researchers

In the light of the reported research findings on the instructional effectiveness of computer and IV simulations, the following conclusions can be drawn:

- All the studies have applied the comparative format;
- none of the studies has compared the effectiveness of computer/IV simulations in teaching different learning tasks;
- none of the studies has compared the effectiveness of computer/IV simulations in teaching a particular learning task in relation to learners' characteristics;
- none of the studies has compared the effectiveness of different types of IV simulations in teaching different learning tasks;
- none of the studies has investigated the effectiveness of different ways (e.g. pair, single) of studying different types of IV simulations in relation to different learning tasks and different learners characteristics;
- little work has been done on the theoretical basis of the design of IV simulations.

As can be seen, there are still criticisms to be made about the way that new technology is evaluated. The striking point seems to be that although the comparative approach has been strongly criticised since the 1950's, evaluators have continued to stick to it. The main reason for the continuity of the comparative approach seems to be that such studies are driven by decision makers' questions, in education, commercial, and military contexts where the main purpose is to prove the effectiveness of the medium introduced by comparing it with the existing one. The question that arises then is: are educators going to continue to evaluate multimedia applications in the same way or are they going to take the criticisms into account? It seems that educators should abandon the task of trying to prove the effectiveness of multimedia simulations in comparison to other technologies and concentrate on exploring the potential of the medium in relation to learning tasks (e.g. problem solving, fact acquisition, concept acquisition, extension of existing knowledge, transfer of knowledge and applications, understanding of procedures and processes, handling of evidence), learners characteristics (e.g. gender, academic achievement, social background).

References

Bates, T., (1981). Towards a Better Research Framework for Evaluating the Effectiveness of Educational Media, BJET, 12, No 3, 215-233.

Clark, R. E., & Salomon, G., (1986). Media in Teaching, in Wittrock, M. (Ed.), Handbook of Research in Teaching, 3rd Edition, London: Collier Macmillan.

Clark, R. E., (1983). Reconsidering Research on Learning From Media, Review of Educational Research, 53, No 4, 445-460.

Clark, R. E., (1984). Clark's Reply to "Clark's "Learning From Media": A Critique", Educational Communication and Technology Journal, 32, No 4, 238-241.
Clark, R. E., (1985). Evidence for Confounding in Computer-Based Instruction Studies: Analyzing the Meta-Analyses, Educational Communication and Technology Journal, 33, No 4, 249-262.
Cohen P., Ebling, B., & Kulik, J. (1981). A Meta Analysis of Outcome Studies of Visual Based Instruction, Educational Communication and Technology Journal, 29, No 1, 26-36.
Cronbach, L. J., & Snow, R. E., (1977). Aptitude and Instructional Method, New York: Irvington.
Moore, D. M., Wilson, L., & Armistead, P., (1986). Media Research: A Graduate Student's Premier, BJET, 17, No 3, 185-193.
Reeves, T. C., (1986). "Research and Evaluation Models For the Study of Interactive Video", Journal of Compater-Based Instruction, 13, No 4, 102-106.
Salomon, G.. & Clark, R. E., (1977). Reexamining the Methodology of Research on Media and Technology in Education, Review of Educational Research, 47, No 1, 99-120.

CHAPTER 5

RESEARCH METHODES AND EVALUATION RESEARCH TECHONOLOGY

687

ADVANCES IN RESEARCH METHODOLOGY AND EVALUATION RESEARCH

Wim van der Linden, Department of Education, University of Twente, Enschede, The Netherlands; Harvey Goldstein, United Kingdom, & Ronald K. Hambleton, USA

Introduction

The papers collected in this section cover a wide range of advances in research methodology and evaluation research. Among the topics dealt with we find such examples as test construction, standardization, and norming, applications of decision theory to education, alternative methodologies to generate and study qualitatieve data in educational research, methods of scaling, evaluation of educational programs, and linear models. The readers are urged to get acquainted with there contents and to experience the ingenious ways in which problems are conceived and solved. The majority of the papers falls into the categories of the methodology of educational assessment, item response theory, and multilevel modelling--and some even successfully explore links between the categories.

Programs of educational assessment usually are large-scale surveys in which data about the progress of curriculum output are collected. The surveys are repeated periodically and show a systematic change of subject matter. Because each survey has to cover a full subject matter area as well as to describe the progress for a national population of examinees, the application of techniques of sampling both items and examinees is inevitable. Typically, the design of programs of educational assessment is longitudinal in character. Policy makers and managers in education want to follow trends in curriculum output over a longer period; therefore results from individual surveys have to be comparable in some sense. Since users can be found at all levels in education, reporting of results has to take place at several levels of aggregation. It is obvious that all requirements constrain the design of assessment programs considerably and that technical ingenuity is required to meet often contradictory requirements. Important methodology is provided both by item response theory and multilevel modelling. Item response theory may be used to scale the test items and develop variables to describe curriculum output and make longitudinal comparisons. Multilevel modelling can be used to design the hierarchical sampling procedure and explain curriculum output by relating the results to factors operational at different levels in the educational system.

Item response theory (IRT) was developed by Frederic Lord and Georg Rasch in the 1950s, but the field had to wait until the 1970s to realize the full potentials of these early developments. Basically, IRT models the probability of a response of an examinee to a test items, parameterizing the properties of the test item and the abilities of the examinee. Since the test item and the examinee are parameterized separately, estimates of the item parameters are automatically adjusted for the abilities of the examinees in the sample, and vice versa. The same feature is used in procedures for adaptive computerized testing, in which the choice of the next item has to match the current estimate of theexaminee's ability and in studies of optimal designs for sampling examinees to calibrate test items. The first models were developed exclusively to deal with dichotomously scored test items. However, in practice test items often have nominal, partial credit, graded, or other non-dichotomous response formats. Also, test items may be represented in testlets, used to measure change or underlying cognitive processes, and response time may be part of their scoring rule. In particular, the dimensionality of the abilities tapped by the test items may be of concern. Modern developments in IRT try covering all these extensions and refinements, and some of the papers in this section show results.

Educational researchers have always been aware of problems involved in research in which the treatments and units of analysis are at different levels of aggregation. At a higher level of generality models for analyzing data from such designs have been available in the statistical literature, but their potential was never fully realized until an interest in multilevel or hierarchical modelling emerged in educational research a decade ago or so. The statistical theory for multilevel modelling can be found in sampling theory for stratified random sampling , the theory of mixture distributions, Bayesian methods, linear models with random effects and regression models with errors in the variables. In fact, one of the oldest multilevel models might be the classical test theory model introduced by C. Spearman, in which the observed variable is assumed to have random error, but at a higher level true scores are assumed to be random as well, as the result of sampling of examinees from a population. Multilevel models share the methodology of careful parametric modelling with other modern applications of statistics in the social sciences. They show that adequate modelling should be based on substantive assumptions about the way reality is organized, that parameters should be introduced for every effect to which the observed results may be liable, and that for each parameter it should be specified whether the effect it models is fixed or random. Only by proceeding in this way can powerful conclusions be drawn.
Just as IRT, a natural application of hierarchical modelling is educational assessment where the models can be used to explain trends and effects in measures of curriculum output. An obvious step is to combine the two inhance the effectiveness and efficiency of current assessment programs.

In this theme the papers are custered into the following categories:
- Adaptive Testing/Optimal Sampling Designs (AT/OSD)
- Applications (AP)
- Applications Rasch Model (ARM)

- Data Collection Management (DCM)
- General Data Analysis (GDA)
- Multi–dimensional (MD)
- Multi–level Applications (MLA)
- Polytomous Items (POIT)

You will find the abbreviations in the first column of the program.

SELF RECORDED AUDIO TAPE: AN ALTERNATIVE QUALITATIVE DATA COLLECTION METHOD

Fred Lockwood, Open University Institute of Educational Technology, United Kingdom

Introduction

Data collection by individual interview, be it face to face or over the telephone, allows great flexibility in the phrasing and sequencing of questions and in the use of neutral prompts, but is limited by the small number of interviews that can be conducted during that window in time when it can be collected. Although held many researchers to be the most appropriate data collection technique its very operation limits the amount of data that can be collected and from which emergent constructs and their interrelationships can be identified. Employing a team of interviewers to increase the data collected is seldom feasible and raises concerns over interviewer reliability. Conducting small group discussions may increase participation but is susceptible to distortion by atypical, and articulate individuals. Furthermore, extending the data collection period incurs other problems including an undue reliance upon the memory and recall of respondants. In contrast collecting qualitative data by self-recorded audio cassette tape can dramatically increase the amount of data available to the researcher; data that is created by the respondent at the most appropriate moment in time.

Studies utilizing the technique

Two separate studies, involving 64 and 40 British Open University students respectively, were conducted in which the dual focus was to explore students' perception and use of Activities in actual teaching material and different qualitative data collection methods. In the first study three data collection methods, involving three separate sub-samples, were employed; the first posed questions in face to face interviews, the second posed them as open ended questionnaire items and the third invited respondants to record their responses on an audio cassette tape. In the second study, in a completely different academic area and again involving several different sub-samples, the three previous data collection methods were supplemented by the use of telephone interviews.

Potential respondants were contacted initial by telephone and invited to participate in the study. Those responding on cassette tape received a copy of the teaching material under investigation, colour coded interview schedules, detailed instructions regarding how to identify themselves and the question to which they were responding and a blank audio cassette tape. The questions were arranged

into those that could be posed before study of the teaching material in question, those that were interleaved into the teaching materials (to be answered as they studied that material) and those that could be answered after their study was complete. Each interview transcript or recording was analysed by three independent judges both to categorize students in terms of their perception and use of Activities and to compare and contrast the quality of data generated by the different data collection methods.

Results of the investigation

The limitations in the use of self-recorded audio tape are obvious - the researcher is unable to establish a rapport with respondants, be sensitive to verbal and non verbal cues, identify and pursue emergent factors and adapt to developments during the interview and types of response. However, analyses revealed that self-recorded tapes made significant contributions to the two studies.

It was evidence from the self-recorded tapes that initially identified one of the constructs within the model that described and explained how students perceive and use Activities in teaching texts; once identified it was confirmed within the data provided by other data collection methods. In subsequent analyses, when the different data collection methods were compared in terms of the successful identification and categorisation of constructs, and in term of levels of response to closed ended questions, the different methods were comparable.

What is more, analyses indicated that relinquishing direct control of the data collection process did not necessarily detract from the evidence that was actually provided, nor did the technique suffer unduly from the inability to curtail or redirect respondants' comments. The technique not only saved time and travel costs, but allowed respondants to decide the most convenient time and place for making a response, eliminated the possibility of leading questions and non verbal cues that may be inadvertently transmitted to the respondent, left the question(s) open to the respondants' interpretation rather than being channelled by the interviewer, avoided any feeling that a constant dialogue be maintained with the respondent controlling the speed and focus of the data collection.

It is worth noting that self recorded tapes proved to be of good technical quality with replies being full and detailed . Indeed, a relaxed conversation pace and conspiratorial tone was a characteristic of self-recorded tapes; the intonation associated with comments, and pauses between them, greatly enhanced the interpretation that was made. Indeed, the solitary conditions under which the tapes were recorded appeared to encourage a degree of openness and frankness that was less apparent in face to face interviews.

Concluding comments

During two separate research studies the use of self recorded audio tape was demonstrated to be straightforward and to present few problems in subsequent analysis. On several indicators considered - length of audio recording, comparison of closed ended responses and identification of constructs, the data collected by self recorded tape compared favourably with that collected by other methods. Further consideration of this form of data collection acknowledged the limitations that self recorded tape imposed but recognized several attractive features worthy of exploitation.

SCORING FUNCTION AND GENERALIZED PARTIAL CREDIT MODEL

Eiji Muraki, Educational Testing Service, & R. Darrell Bock, University of Chicago, USA

The Generalized Partial Credit model (Muraki, 1991) for the j–th item and the k–th categorical response among the total of m_j categorical responses is expressed by

$$P_{jk}(\theta) = \frac{\exp\left[\sum_{v=1}^{k} Z_{jv}(\theta)\right]}{\sum_{c=1}^{m_j} \exp\left[\sum_{v=1}^{c} Z_{jv}(\theta)\right]} \qquad (1)$$

where

$$Z_{jv}(\theta) = a_j(\theta - b_j + d_v) \qquad (2)$$

and a_j is a slope parameter, b_j is a location parameter, and d_v is a threshold parameter. The sum of $Z_{jv}(\theta)$ can be written as

$$Z^*_{jk}(\theta) = \sum_{v=1}^{k} Z_{jv}(\theta) = a_j\left[(k(\theta - b_j) + \sum_{v=1}^{k} d_v\right] \qquad (3)$$

The model above can be rewritten as

$$Z^*_{jk}(\theta) = a_j[T_k(\theta - b_j) + K_k] \qquad (4)$$

Andrich (1978) calls T_k and K_k the scoring function and the category coefficient, respectively. For the Partial Credit model, the scoring function T_k is a linear integer scoring function, that is, $\underline{T}=(1, 2, 3, ..., mj)$ where m_j is the number of categories of item j.

The Partial Credit model becomes the model for ordered categorical responses only when the scoring function is increasing, that is, $T_{k+1} > T_k$, for any k and $a_j>0$. Thus, if we assign an identical scoring to two response categories, we can construct the Partial Credit model with partially non–ordered categorical responses. If the scoring, T_k and $T_{k'}$, are identical, the odds ratio, $R_{j|k'k}$, of these categorical responses, $P_{jk'}(\theta)$ and $P_{jk}(\theta)$, is independent of the latent trait, θ. The odds ratio becomes the function of the difference of the categorical coefficients, that is,

$$R_{j,k'|k} = \exp[a_j(K_{k'}-K_k)] \qquad (5)$$

The odds ratio is constant along the scale and these item category characteristic curves (ICCC) never intersect. The Partial Credit model with partially identical scoring functions is called the Partially Unordered Partial Credit model (PUPCM). The NAEP (National Assessment of Educational Progress) scoring rubric counts omits as wrong responses. It is frequently observed that the open ended question induces high frequencies of omitted responses, especially for minorities. However, after carefullyanalyzing item responses patterns among ethnic groups, it was found that ethnic differences in responding can be explained to a great extent by ability differences. Thus, it seems reasonable to parameterize omitted responses for the Partial Credit model as a function of θ. If we use the Partial Credit model to analyze the item response data with high occurrence of omitted responses, we either exclude the omitted responses or make some assumptions about the omitted responses. If we assume that the omitted response is the response category lower than the lowest categorical response, usually the wrong response, we apply the Partial Credit model directly to the analysis of the response data. If we do not assume the inherent order between the wrong response and the omitted response, we analyze the responses by using the PUPCM.
In this paper, the development of the PUPCM will be described and the application of this model to the NAEP data will be examined. The comparison of this model with Bock's nominal response model (Bock, 1972) will be also discussed.

References

Andrich, D. (1978). A rating formulation for ordered response categories. Psychometrika, 43, 561–573.

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. Psychometrika, 37, 29–51.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. Applied Psychological Measurement (in press).

CONSTRUCTING MEASUREMENT VARIABLES WITH THE RASCH PARTIAL CREDIT MODEL

G.N. Masters, The Australian Council of Educational Research, Australia

This paper considers issues in using models for ordered response categories-and in particular the Rasch partial credit model- to mark out and describe measurement variables. In the case of dichotomous items, the Rasch model provides a single parameter estimate (treshold) for an item enabling each dichotomous item to be mapped to a single location on a variable and used to construct an interpretation of that variable (a process sometimes described a 'variable definition' or 'behavioral anchoring'). In the case of items scored in mord than two ordered categories, the Rasch model provides more than one parameter for each item. This paper will show how for the purposes of variable definition, the parameters of the partial credit model can be usefully used to construct an alternative set of parameters that have the same meaning as tresholds (or category boundaries) in the models of Thurstone and Samejima. The use of Thurstone tresholds to construct and describe a scale for reporting the results of an Australian statewide testing program will be described.

A PERSON FIT FUNCTION SENSITIVE TO ITEM RELIABILITY AND DISTANCE BETWEEN ITEMS

Rob R. Meijer, & Klaas Sijtsma, Vrije Universiteit, Amsterdam, The Netherlands

In person fit analysis it is investigated whether a person exhibits a pattern of item scores that deviates from patterns that are predicted by a measurement model or from the patterns of a majority of persons in the population to which the person belongs. For persons that are detected as aberrant the total score does not adequately reflect their ability. Further research is needed before any firm conclusion can be drawn about test performance.

Recently, many methods for detecting aberrant response patterns have been proposed using parametric item response theory (IRT) models (e.g. Levine & Drasgow, 1982; Molenaar & Hoijtink, 1990). Also in the context of nonparametric models and in the context of group-based indices several methods are proposed (Sato, 1975, Harnisch & Linn, 1981).

This paper concentrates on the latter approach. Two groups of indices can be distinguished. One group is based on the consistency of the response pattern with the group based difficulty order of the items (e.g. the caution index, Sato, 1975). Deviance is expressed using a covariance or correlation measure. Another group of indices uses the group based difficulty order to determine the number of Guttman errors in an observed item score pattern (e.g. the norm conformity index, Tatsuoka & Tatsuoka, 1982). This last restriction predicts that given a person's total score x on k items, the x easiest items in the group are correctly answered and the k-x most difficult items are incorrectly answered.

In general, each person fit method is particularly sensitive to score patterns with positive answers on difficult items and negative answers on easy items. Furthermore, Molenaar and Hoijtink (1990) note that all person fit indices use the spread of the item difficulties ($\hat{I}$ from IRT or π from classical test theory). Zero variance of the $\hat{I}$s means that person fit indices are useless and the larger the variance is, the more favorable the conditions for person fit are. Additionally, Meijer, Sijtsma and Molenaar (1991) show that the evaluation of item score patterns is highly affected by the reliability of the items on which the analysis is based: The higher the item reliability, the more favorable the conditions for person fit. Although it is recognized that person fit indices are sensitive to the spread of the item difficulties and the reliability of the items, no person fit measure exists that explicitly uses these two sources of information to classify a response pattern as either aberrant or normal. In this paper such a measure is proposed.

To obtain a person fit index that is sensitive to item reliability it is necessary to obtain reliability estimates. Mokken (1971) proposes two methods to approximate reliability of test scores. A third method (method MS) was proposed by Sijtsma and Molenaar (1987). These three methods were proposed in the context of test reliability, however. Meijer et al. (1991) investigated properties of these methods applied to individual items. By means of a Monte Carlo study, it was shown that method MS was always less biased and the standard deviation across samples was almost always smaller compared to the other two methods. Meijer et al. concluded that method MS performed at least as good as the other two methods and should therefore be prefered.

In the person fit function to be proposed here, the Guttman errors are weighted by a function of the individual item reliabilities and the distance between the items. Let us denote the proportion correct score on the items g and h by π_g and π_h and the item reliablity on these items by $\rho_g$and ρ_h. Let us further define the distance between two items g and h as the difference between the proportions correct score. The distance is denoted β_{gh}. Given that the items are ordered according to decreasing Î, that is increasing proportion correct on the items, $\beta_{gh} = \pi_h - \pi_g (g=1, ...,k-1; h=g+1,...,k)$. The weights are given by:

$$W_{gh} = \min(\rho_g,\rho_h)\beta gh \qquad (1)$$

where $\min(\rho_g,\rho_h)$ denotes the smallest item reliability value of the items g and h. Note that $W_{gh} = 0$ if either ρ_g or ρ_h or both equal zero. Also $W_{gh} = 0$ if the distance between the items equals zero. The maximum value is obtained if the reliability of both items equals one and the distance between the items is at its maximum given the total score and the number of items.

To obtain a person fit measure that is easy to interpret and is independent on the number of items, it is desirable to norm these weights. At the moment it is investigated which definition of weights yields the most desirable results. By means of simulated and empirical data the use of several alternative functions is illustrated.

References

Harnisch, D.L., & Linn, R.L. (1981). Analysis of item response patterns: Questionable test data and dissimular curriculum practices. Journal of Educational Measurement, 18, 133-146.

Levine, M.V., & Drasgow, F. (1982). Appropriateness measurement: Review, critique and validating studies. British Journal of Mathematical and Statistical Psychology, 35, 42-56.

Meijer, R.R., Sijtsma, K. & Molenaar, I.W. (1991). Item reliability and person fit (submitted for publication).

Mokken, R.J. (1971). A Theory and Procedure of Scale Analysis. The Hague: Mouton/New York, Berlin: de Gruyter.

Molenaar, I.W., & Hoijtink, H. (1990). The many null distributions of person fit indices. Psychometrika, 55, 75-106.

Sato, T. (1975). The Construction and Interpretation of S-P Tables. Tokyo: Meiji Tosho (In Japanese).

Sijtsma, K., & Molenaar, I.W. (1987). Reliability of test scores in nonparametric item response theory. Psychometrika, 52, 79-97.
Tatsuoka, K. K. & Tatsuoka, M.M. (1982). Detection of aberrant response patterns. Journal of Educational Measurement, 20, 221-230.

RASCH ANALYSIS AS AN INSTRUMENT FOR MEASUREMENT IN THE CASE OF MISSING DATA

P. Alvarez, Faculty of Economics, University of Extremadura, Badajoz, Spain

Rasch model is one of the most representative model, that explain the process when data are given as a result of the interaction between the person and the item. We are trying to analyse the role of Mathematics, considered as a service subject, is fulfilling its goal when a lot of data are missing.
The probability that one student n pass a subject i in a dichotomous case is given by

$$Pr[X_{ni}=1] = \frac{e^{(\beta_n-\delta_i)}}{1+e^{(\beta_n-\delta_i)}}$$

the probability to fail is

$$Pr[X_{ni}=0] = \frac{1}{1+e^{(\beta_n-\delta_i)}}$$

where β_n is the student ability parameter and ∂_i is the subject difficulty parameter.

An important characteristic of the Rasch model is that the probability can be only expresed either in terms of the ability or difficulty. An extension of the dichotomous case is when the items have more than one response categories; that is the case of The Extended Logistic Model (ELM). In the present case the categories are 1,2,3,4,5,6,7,8,9 which are the marks of the exam results for the students.

Data are the exam results of the 227 second year Mechanical and Electrical engineer students in 12 differents subjects. Not all students attended all the exams, since not all of them were enrolled in all the subjects. From that twelve subjects four are commom subjects for both students, other four subjects are only for Mechanical engineering students and another four are for Electrical engineering students. So the range of subjects for each student vary from one to eight.

The convergence of parameters have been reached after 40 iterations. Index for fitness is all right, so the data fits well the Rasch model.

In the following table for the "expected score" appear the item characteristic curves for different categories, the scale of difficulty is located at the bottom and all the subjects are on the left hand side. Difficult subjects are located at the top of the table and the easiest ones at the bottom. Then, for a given degree of difficulty, say 53.00, one student who has a mark of 5 in Electronics 1, he would expect to get a mark of 5.5 in Electronis 2, 5.5 in Thermodinamics, 5.5 in Mathematics, 6 in Mechanics Technology, 7 in Strength Materials, and 8 in English.

```
EXPECTED SCORE  (BETWEEN "0" AND "1" IS "0", ETC.)
 NUM COURSE     20.0  26.0  32.0  38.0  44.0  50.0  56.0  62.0  68.0  74.0 80.0
                |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
   8 E ELCTRNCS 1 1           2           3     4  5  6     7     8        9  9
   7 E CIRC.THRY  1           2           3     4     5  6     7     8     9  9
   9 E ELCTRNCS 2 1           2           3     4  5  6     7     8     9  9
   2 C THRMODYNMC 1           2           3     4  5  6     7     8     9  9
   1 C MATH       1           2           3     4  5  6     7     8     9  9
   3 M MECH.TECH  1        2           3     4  5  6     7     8        9  9
   5 E ELEC.DRAW. 1        2           3     4  5  6     7     8     9     9
   6 C MAT.SCI.   1        2           3     4  5  6     7     8     9     9
  11 M TECH.MECH. 1        2        3     4  5  6     7     8     9        9
  10 M STRNGTH/MT 1     2           3     4  5  6     7     8     9        9
  12 M MECH.DRAW. 1     2           3     4  5  6     7     8     9        9
   4 C ENGLISH    12          3     4  5  6     7     8     9              9
                |-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
               20.0  26.0  32.0  38.0  44.0  50.0  56.0  62.0  68.0  74.0 80.0

                                                1 113221  1
PUPILS          1                  22 1 7 7354731196979573161   1
                                      0     3     M     3     0
```

The most difficult subject is Electronics 1, since has the higher calibration, and the easiest one is English, as it is showed in the next table.

COURSE STATISTICS -- CALIBRN ORDER

NUM	COUNT	SAMPLE	CALIBRTN	ERROR	MNSQ	INFIT	MNSQ	OUTFT	PTBIS	NAME	M	G
8	145	31	54.8	.7	1.20	.8	1.23	.9	.37	E ELCTRNCS 1		
7	149	31	53.4	.7	.64	-1.5	.70	-1.2	.33	E CIRC.THRY		
9	221	46	53.3	.6	1.78	4.0	1.81	4.0	.24	E ELCTRNCS 2		
2	383	80	53.3	.5	1.12	.8	1.12	.8	.07	C THRMODYNMCS		
1	326	69	53.1	.5	.88	-.8	.83	-1.1	.12	C MATH		
3	395	74	51.6	.5	.67	-2.1	.72	-1.7	.32	M MECH.TECH		
5	260	50	50.2	.6	.81	-1.0	.85	-.8	.38	E ELEC.DRAW.		
6	474	88	49.7	.5	.75	-1.6	.73	-1.8	.20	C MAT.SCI.		
11	115	19	48.3	1.0	.36	-1.9	.39	-1.6	.57	M TECH.MECH.		
10	106	16	45.5	1.2	.88	-.3	.89	-.3	.21	M STRNGTH/MTRLS		
12	124	19	45.1	1.1	.91	-.3	1.08	.2	.52	M MECH.DRAW.		
4	800	112	41.7	.5	1.11	.8	1.14	1.1	.04	C ENGLISH		

Codes has been written for each subject. The letter E means that is a subject for Electrical engineers, M for Mechanical, and C is a common subject. According to the calibration order output, the subjects refering to E have bigger calibrations than subjects C and M.

Paying attention to Mathematics, we know that if the goal for Mathematics is being considered as a service subject, then its difficulty should not be bigger than the subject for which is providing such service. In our case we see that Mathematics calibration is below than Electrical engineering subjects and above Mechanical

engineering subjects. Then mathematics plays its role well for the Electrical engineering students. On the other hand, Mathematics is too difficult for the Mechanical engineering students. Hence Mathematics should not be a common subject, its content should be revised. Therefore it should be splitted in two: Mathematics for Electrical engineering students and Mathematics for Mechanical engineering students.

To the question of what speciality is more difficult, it can be noticed according to the table that Electrical speciality is more difficult than the Mechanical one. Another risen question is what are the things that make to Electronics I so difficult? Is it the lack of basic knowledge from the previous courses?, etc.

The following table shows only part of the list of students refering to their abilities. Num is the student number; count is their score in the exams; test is the number of exams taken; measure is the student ability, etc.

PUPIL STATISTICS -- MEASURE ORDER

NUM	COUNT	TEST	MEASURE	ERROR	INSQ	INFIT	OUTSQ	OUTFT	PTBIS	NAME	M G
199	8	1	67.9	5.7	.00	-.7	.00	-.7	.00	1322	
112	22	3	63.7	2.9	1.79	1.0	1.77	1.0	-.05	1495	
151	7	1	61.9	5.0	.00	-.7	.00	-.7	.00	1630	
213	7	1	61.9	5.0	.00	-.7	.00	-.7	.00	1244	
228	7	1	61.9	5.0	.00	-.7	.00	-.7	.00	1713	
65	30	4	61.8	2.7	.85	-.2	.79	-.3	.76	1759	
175	7	1	61.7	5.0	.00	-.7	.00	-.7	.00	1270	
226	7	1	61.7	5.0	.00	-.7	.00	-.7	.00	1204	
132	14	2	61.1	3.5	.03	-1.0	.03	-1.0	.00	1508	
230	7	1	60.2	5.0	.00	-.7	.00	-.7	.00	1777	
9	58	8	59.8	1.8	1.24	.5	1.34	.7	.14	1665	

It can be seen the usefulness of Rasch model, considered as an instrument for measurement, providing criteria to make decisions in this special case.

This work is only a sample of a bigger project that try to calibrate difficulties for subjects from the same careers in different EEC countries, according to students marks.

Reference

Andrich, A. (1988). Rasch Models for Measurements, Murdoch University, Sage.

Wright, B.D., & Stone, M.H. (1970). Best Test of Design". University of Chicago, Mesa press.

Wright, B.D., & Masters, G.N. (1982). Rating Scale Analysis, University of Chicago, Mesa press.

EVALUATING THE DEVELOPMENT OF CRITICAL THINKING IN UNDERGRADUATE NURSES USING RASCH MEASUREMENT

B.E. Sheridan, Edith Cowan University, Western Austria

Introduction

This project is designed to evaluate the effectiveness of a course of study in critical thinking in an undergraduate nursing program. The logic of the measurement and analysis techniques involved will provide the emphasis of this paper rather than what constitutes critical thinking and how it is conceptualized and presented to undergraduate students.

To investigate the problem, the familiar repeated measures strategy in association with a pretest/posttest design provides a logical starting point. Instead of using raw scores to represent the unit of analysis, as is often the case in studies of this kind, the Extended Logistic Model (ELM) of Rasch (Andrich, 1985, 1988) will be employed to provide interval measures of known precision. Besides assessing the dimensionally of the instrument, advantage can be taken of the ability of this model to account for extended response categories of differing sizes within the scoring schedule of the instrument. data collected from the experimental and control groups at the pretest and posttest occasions are first assessed on an individual basis for the presence of interaction effects that will bias the precision of the measures involved. Once the validity of the instrument has been established, the data from all occasions can then be pooled to provide more precise measurement estimates than is otherwise possible from the individual occasions. In this way, a claim of true commensurate measures can be established.

Method

Subjects

All newly enrolled students in Semester 1, 1992, for the Bachelor of Nursing degree (N = 200) at Edith Cowan University in Western Australia (total enrollments over 16,000) are involved in the project. Approximately equal numbers of students (N = 50) are allocated, at time of enrollment, to four different elective course units in as near a random manner as possible. The course unit, Critical Thinking, is designated the experimental group Yhile students enrolled in the three other elective units -- Introduction to Psychology, Introduction to Sociology, and Intercultural Perspectives -- provide a series of control groups.

702

Assessment

The measuring instrument has been created especially for this project. This comprises 12 separate items, most of which employ an extended response format whereby student responses are assessed and coded according to a specified schedule. Answers comprise written explanations -- maximum of four lines -- of the logic, or otherwise, of a series of situations describing scenes that nurses will most likely encounter within the execution of their duties.

Data analysis

Data is collected from students, both prior to the commencement of the teaching sessions (the pretest occasion) and upon completion of the four Semester 1 units of instruction identified earlier (the posttest occasion), and using the same instrument throughout. These responses will be coded by lecturers involved with educational philosophy and separate data files compiled. Anonymity and linking of data is accomplished using numbers allocated to students.

Results

An item analysis on both the pretest data (available February/early March) and the posttest data (available June/early July) to check for interaction effects due to bias in the data comprises stage one of the overall analysis for the project. then sound measurement has been established, the data is pooled and precise estimates obtained for the commensurate measures for the next stage of the analysis. The logic of this latter step in the item analysis is that responses for each student will appear as two separate data records within the data file: one for the pretest and one for the posttest occasion. Thus, if the instrument is a true measure critical thinking, then the response-pairs for each student will represent, in effect, two levels of ability for the experimental group (if development of critical thinking has occurred during the semester) and possibly, even for the control groups (where trend effects, such as normal maturation, may also be present).

Once the measure is established for the instrument, individual ability estimates are determined and allocated to the appropriate cells of the pretest/posttest design. In this project, the dependent variables approach advocated by Bock (1975) is preferred to that of the two factor ANOVA model of Winer (1972).

Conclusion

Because of the time line for this project, no conclusions regarding the data are possible for this summary statement. discussion on the development of a measure of critical thinking will be available for the pretest data collected in February. This would involve the analysis of multiple category data using a measurement model capable of providing a measure of the trait involved.

Overall, the basic structure of the instrument is flexible and could be adapted to involve other discipline areas. Data collected from such extensions could then be compared with that obtained for this project, thus providing further insights into the validity of the instrument as a measure of critical thinking.

NON-PARAMETRIC MODELS FOR MULTICATEGORY: ITEM RESPONSE MODELS REVIEWED AND EXTENDED

I.W.Molenaar, Statistics and Measurement FPPSW, University of Groningen, Groningen, The Netherlands

Introduction

For tests measuring abilities, achievements or attitudes, item response theory (IRT) models answer probabilities per item per subject. This is in many respects superior to classical test theory which is a model for the total score on the test (see e.g. Hambleton & Swaminathan, 1985). Most IRT models assume dichotomous scoring, although the original score per item is often more informative than a dichotomy. The taxonomy by Thissen & Steinberg (1986) shows how both ordered and nominal answer categories may find a place within the IRT model rather than be recoded into 1 (= correct/agree) and 0 (= incorrect/disagree) before further data analysis takes place. Valuable information may get lost during dichotomization; Jansen and Roskam (1986) even show that the Rasch model can not be valid both before and after dichotomizing graded responses, or simultaneously for different dichotomizations.

In applied educational research, however, as well as in applied work in other social and behavioral sciences, the tradition of using dichotomized scores still persists. This may be partly due to the complicated nature of multicategory IRT models, and to the less widespread availability of software for their parameter estimation and fit analysis. The present paper discusses a class of non-parametric IRT models based on simple frequency counts, in which item categories and subject positions can be ordered on the latent trait if the item response curves fulfill some simple order conditions.

Method

Postponing both point items and nominal categories until the last section, let us assume that a set of k given items with m+1 ordered categories each have been presented to a random sample of n subjects. Categories like wrong/partially right/correct or strongly disagree/disagree/agree/strongly agree will be scored as 0,1,...,m with m=2 and m=3 respectively. Following Masters (1982), each item will be decomposed into m "item steps" here defined as "scoring $\geq$j rather than <j" for j=1,2,...,m. It will be assumed that the item response function (IRF), which is the probability of a positive answer to such an item step, is an increasing function of the latent trait value of the subject, and that local independence holds between

steps belonging to different items : given someone's position on the latent trait measured by the items, answers to different items are statistically independent. As is shown in Molenaar (1982), this extension of the dichotomous non-parametric model of Mokken & Lewis (1982) allows ordering the subjects by their summed Likert score on all items, and ordering the item steps by their popularity in the sample. The computer program MSP (Debets & Brouwer, 1989) allows the search for scales conformal to this model from a larger item pool as well as the fit test of existing scales. It is reasonable to require that the distribution of the latent trait givena value of the sum score is stochastically increasing with this sum score; for dichotomous items with increasing IRF's this was proven by Grayson (1988), but for multicategory items an additional technical assumption is required. Moreover, Molenaar (1991) has proposed an improved definition of the Loevinger H coefficient for the multicategory case by which it keeps its interpretation as the correlation between two item scores divided by the maximal correlation given the marginals. This will be incorporated in the next release of MSP.
With the additional assumption that the IRF's are non–intersecting, the Mokken model allows to assess the reliability of the person scores with a method that is superior to Cronbach's alfa (Sijtsma & Molenaar, 1987).

Results
The scaling method outlined above has been successfully applied to scales measuring a variety of latent traits such as arithmetic ability, health complaints after computer screen work, psychiatric disability, Piagetian seriation, trustworthiness of nations and sexual arousal capability. In most applications, the joint scaling of persons and item steps on the latent dimension produces not only an estimated ordering of the persons on the latent trait; via the order of the item steps additional insight in the structure of the scale also comes within reach, and expectations about the role of specific items can be empirically tested.

Discussion
The Mokken model is widely applicable, as it only assumes some mild order restrictions on the IRF's. It is also easy to explain and fast to apply : calculations are simple and there is no need for iterative parameter estimation methods as used in logistic or normal ogive IRT models. The step from two to more ordered answer categories is less difficult than for parametric IRT. The use of Likert scores rather than dichotomized scores leads to more refined scaling of the persons on the latent trait.
Of course there is a price to pay for these advantages : one obtains only order information, no metric information on the positions of persons and item steps. This is a common complaint about nonparametric methods. If a stronger parametric model fits the data, and if metric information is important, one should probably not use the Mokken model, or perhaps pass to the stronger model after a preliminary Mokken analysis.
Just like with parametric IRT models, the model assumptions are formulated in terms of the unknown IRF's and latent trait values. Therefore they can only be indirectly checked from their manifest consequences. Currently, some new fit statistics for the multicategory Mokken model are implemented in the MSP program. Their usefulness in detecting model violations without deleting model

conformal items was established in both real and simulated data sets, but the minimal assumptions of the model make adequate fit assessment more difficult than in e.g. the Rasch model. On the other hand, in many areas it is very difficult to find scales meeting all requirements of the latter, whereas good Mokken scales are easier to obtain.

It is an intriguing question to what extent similar methods could be used for nominal answer categories. For multiple choice items in an achievement test, the conventional right/wrong recoding usually involves little loss of information (but see Thissen & Steinberg, 1984). With attitude measurement, it may happen that one of the categories has a unimodal IRF and no distinction between the other nominal categories is important.

There is an analog to Mokken scaling for dichotomous point items (e.g. in preference judgements), where the probability of a high score on an item decreases with the distance of the person position from the item position (unimodal trace lines), see Van Schuur (1984) and Post (1992). Finally, an extension by Van Schuur (1991) is in progress for the case of multicategory point items (e.g. strongly preferred, possibly perferred, not at all).

References

Debets, P. & Brouwer, E. (1989). User's Manual MSP. Groningen : iec ProGAMMA.

Grayson, D.A. (1988). Two–group classification in latent trait theory : scores with monotone likelihood ratio. Psychometrika 53, 383–392.

Hambleton, R.K. & Swaminathan, H. (1985). Item Response Theory. Boston : Kluwer–Nijhoff.

Jansen, P.G.W. & Roskam, E.E. (1986). Latent trait models and dichotomization of graded responses. Psychometrika 51, 69–91.

Masters, G.N. (1982). A Rasch model for partial credit scoring. Psychometrika 47, 149–174.

Mokken, R.J. & Lewis, C. (1982). A nonparametric approach to the analysis of dichotomous item responses. Applied Psychological Measurement 6, 417–430. Molenaar, I.W. (1982). Mokken Scaling Revisited. Kwantitatieve Methoden 3, 145–164.

Molenaar, I.W. (1991). A weighted Loevinger H–coefficient extending Mokken scaling to multicategory items. Kwantitatieve Methoden 37, 97–117.

Post, W.J. (1992). Nonparametric unfolding models, a latent structure approach. Leiden : DSWO Press.

Sijtsma, K. & Molenaar, I.W. (1987). Reliability of test scores in nonparametric item response theory. Psychometrika 52, 79–97.

Thissen, D. & Steinberg, L. (1984). A response model for choice items. Psychometrika 49, 501–519.

Thissen, D. & Steinberg, L. (1986). A taxonomy of item response models. Psychometrika 51, 567–577.

Van Schuur, W.H. (1984). Structure in political beliefs.a new unfolding model with aopplication to European party activists. Amsterdam : C.T.Press.

Van Schuur, W.H. (1991). Nonparametric unidimensional unfolding for multicategory data. Congress paper 1991, now under review for Political Analysis.

706

A MULTIDIMENSIONAL POLYTOMOUS LATENT TRAIT MODEL THAT EXPLAINS THE EXAMINEES' CHOICE OF PARTIALLY CORRECT ALTERNATIVES IN MULTIPLE CHOICE ITEMS

Henk Kelderman, University of Twente, Enschede, The Netherlands

In explaining the responses on a multiple choice items one may wish to specify a model that explains the cognitive processes that underlie the choice of a particular distractor. To do this, the analyst may want to specify a response function for each response alternative. Each response function specifies which latent traits may be responsible and how strong the relation is. Each item may have a different number of response alternatives. An extension of Rasch' (1961) multidimensional model that gives the analyst this flexibility in specifying the relation between item responses and latent traits is the Multidimensional Polytomous Latent Trait (MPLT) model (Kelderman, 1988).

Let B_{qjx} be a positive integer valued weight of response x of item j with respect to the q^{th} latent trait. If $B_{qjx} \neq 0$, it means that the item response denoted by the pair (j,x) depends on latent trait q and if $B_{qjx} = 0$ it does not. In addition, if the weight is larger than one, it means that the response involves more than one application of the latent trait. Furthermore let θ_i be a parameter describing the ability of person i and let δ_{jx} be a difficulty parameter of response x on item j. The MPLT now describes the probability π_{ijx} of a response x of person i to item j as:

$$\ln \pi_{ijx} = \mu_j + B_{1jx}\,\theta_{i1} + \ldots + B_{sjx}\,\theta_s - \delta_{jx},$$

j = 1,...,k; x = 0, ..., r_j. As described by Kelderman (1988) additional constraints must be imposed on the parameters to obtain a unique set of parameter estimates. Furthermore he (Kelderman, 1988, 1992 in press) describes an a method to obtain maximum likelihood estimates of the parameters and to test the fit of the model to the observed data.

An Example

The American Society of Clinical Pathologists (ASCP) produces tests for the certification of medical personnel. The data we analyze here are the responses of 333 examinees to 9 four-choice items measuring the ability to perform medical

laboratory tests. The items are calibrated under a Rasch model so that the sum of the correct answers contains all information about the subject's ability θ_i available in the data.

There are, however, reasons to believe that this single ability parameter might not be sufficient to explain the subjects' behavior on the tests. In particular, it was hypothesized that several different cognitive processes are involved in making the items and that even the incorrect responses might be chosen on the basis of partial execution of these processes. The correct response would then be chosen if all processes where successfully executed and orchestrated.

Table 1 gives the MPLT weights (Bqjx) given by ASCP content experts about three cognitive processes that are possibly involved in choosing the items alternatives. For example, in Item 3 the correct answer b involves the application of knowledge as well as two computations, whereas in the incorrect answer c one calculation is missing.

To investigate whether each activity corresponds to a separate cognitive process, we specify two models.

- (i) A model with one latent trait (θ_{i4}) where the B-weights correspond to the correct/incorrect category IV.
- (ii) A model where all entries in are B-weights, where each activity is a separate latent trait (with subject parameters θ_{i1}, θ_{1i}, θ_{i3}, θ_{i4}).

Table 1

Specification of Cognitive Processes Involved in the Item Responses of the ACSP Medical Laboratory Test.

Item	I Applies Knowledge				II Calculates				III Correlates Data				IV Correct			
	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d
1			2	1											1	
2	1		2												1	
3		1	1			2	1							1		
4	2	1	1		1								1			
5	1		1	1			1								1	
6	1	1								1				1		
7	1	1	2	1							1				1	
8	1	1			2	1			1				1			
9	1		1		1		1				1				1	

N = 3370, Non response = 39

708

The item difficulty parameters of both models where estimated with the LOGIMO program (Kelderman & Steen, 1988, Kelderman, 1992). The log-likelihood (and number of independent parameters) of model (i) is -1782 (36) and of model (ii) is -1104 (446). Akaike's (1977) information criterion then becomes 1854 and 1996 respectively. Since the latter is larger, the conclusion must be that the cognitive processes do not explain the structure of the data beyond the simple correct model. In Table 2 are the item difficulty parameters of model (i).

Table 2

Parameter Estimates of Model (i).

Response	Item								
	1	2	3	4	5	6	7	8	9
a	-0.54	-0.88	-0.37	0.01*	-0.14	1.13	-0.30	-0.46*	0.97
b	-0.95	0.02	-0.66*	0.07	-0.93	-0.07*	0.60	1.05	-1.29
c	0.92*	0.50*	1.00	-0.48	-0.47*	-0.58	0.21*	-0.16	0.05*
d	1.50	0.85	-0.63	0.41	1.07	-0.54	-0.30	-0.90	0.31

* Correct Response

Discussion

In this paper a general multidimensional Rasch model for polytomously scored items is introduced. The analysis of the ASCP data shows that multidimensionality can be modelled quite flexibly on the item response level. It is shown that multidimensionality is not present in these data and that a unidimensional model suffices to describe the data.

References

Akaike, H. (1977). On entropy maximization principle. In P.R. Krisschnaiah (Ed.), Applications of statistics (pp. 27-41). Amsterdam: North Holland.

Kelderman, H. (1988, in review). Loglinear multidimensional IRT models for polytomously scored items. (ERIC document Reproduction Service No. ED 308 238)).

Kelderman, H. (1992, in press). Computing maximum likelihood estimates of loglinear IRT models from marginal sums with special attention to loglinear Item Response Theory, Psychometrika,.

Kelderman, H., & Steen, R. (1988). LOGIMO: Loglinear IRT modeling. [Program Manual]. Enschede, The Netherlands: University of Twente.

Lord, F.M., & Novick, M.R. (1968). Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley.

Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Copenhagen: Paedagogiske Institut.

Rasch, G. (1961). On the meaning of measurement in psychology. In J. Neyman (Ed.), Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 5. Berkeley: University of California Press.

Rasch, G. (1977). On specific objectivity: An attempt at formalizing the request for generality and validity of scientific statements. Danish Year-book of Philosophy, 17, 58-94.

Rasch, G. (1980). Probabilistic models for some intelligence and attainment tests. Chicago: The University of Chicago Press.

710

SCALOGRAM ANALYSIS ON POLYTOMOUS ITEMS

P.V. Zysno, Germany

During the forties Guttman developed the scalogram analysis. Where adequate reactions occur this method yields rank orderings of people and items with respect to the degrees of the defined property. If the lines and columns of the data matrix are rearranged in the appropriate manner. the reaction patterns line up in the form of a regular parallelogram. Torgerson (1958) showed that the perfect regularity has to be weakened in the case of polytomous items. This important modification. however, found little response, probably due to the fact that the scaling problem remained technically unsolved. as the original approach via the marginal sums proved to be unsatisfactory for polytomous items.

Further development continued in two main lines both restricted to dichotomous items. The first was concerned with appropriate tests of goodness of fit, the second directed to the axiomatic foundation which was successfully carried out in 1969 by Ducamp and Falmagne. A scaling algorithm for polytomous items is still missing. On technical grounds substantially polytomous items have to be divided either into two sections or into a set of dichotomous items.

Here a generalized method of scalogram analysis is presented, which processes items with any number of response categories. All of Guttman's assumptions are maintained. On the basis of a set theoretical rationale two different calculation specifications are developed in order to determine on the one hand the rank order of people. and on the other that of category boundaries. The two orderings are interlocked in a composite system. The individual patterns can be visualized in a semi-regular parallelogram.

The model test investigates each 2x2-field as the smallest structural unit which encompasses the combined dominance relationship of subjects and category boundaries. An error is identified if the dominance relation in a pair of boundaries is regarded as contrary by two individuals. The chosen error definition includes the violation of the bitransitivity axiom for dichotomous items. The fit of the data can be checked by a test of significance.

The usefulness of the technique and its mode of operation is shown by some examples of application.

MEASUREMENT ADVANCES TO ADDRESS EDUCATIONAL POLICY QUESTIONS

Ronald K. Hambleton, University of Massachusetts at Amhers, USA

In addressing UNESCO at their general conference in Paris in November of 1991, IEA Chairman Tjeerd Plomp advanced a number of important educational questions:

- What will the world of schools be like for children in the year 2000?
- How will we know what our children are learning and how they are feeling about what they learn?
- How do we know whether our children receive similar educational opportunities and learn the essential basics in comparison with children from similar or neighboring countries?

Educational policy makers want answers to these and other related questions and, in the United States, for example, research funds are often available to educational researchers and measurement specialists who are willing to become involved in policy-related research.

The primary purpose of this paper is to focus attention on four advances in measurement that appear to enhance the validity of test results, and thereby contribute to the usefulness of test data for policy-related decision making: item formats, standard-setting, measurement models, and translating tests.

Item Formats

For more than 50 years in the United States, the multiple-choice item format has been popular in national achievement, aptitude, and credentialing tests, and classroom tests as well. However, in recent times, as policy-makers have begun to take test results more seriously for monitoring educational progress, criticisms from many quarters have followed. Though some of the criticisms appear to be incorrect (see, for example, Hambleton & Murphy, 1992), nevertheless, pressure from critics and those affected by test results has led to a renewed interest in performance testing (called "authentic testing" by the critics), and in the development of new item formats to facilitate machine-scoring and/or the assessment of higher-order thinking skills.

As for new item formats in wide-scale use, the list includes "standardized patients", computer-based problem solving, audio-visual context setting, performance, short answer, essay, portfolio, multiple true-false and problem sets. These and other new formats are being carefully developed and evaluated for their applicability in national testing programs. Criteria for evaluation besides test score reliability and validity include ease of use and costs. Currently, many of these new formats appear to deliver acceptable levels of reliability and validity but the costs of development are unacceptably high and they are difficult to administer and score. Policy makers now have a major dilemma - they can often improve the validity of tests via the use of new item formats, but the financial costs associated with any of these new formats

can be substantial. On the other hand, items which are in one of the standard item formats such as multiple-choice are less expensive to develop and to score but critics feel that this type of item reduces test validity and negatively impacts on classroom instruction.

Standard-Setting

For the past 20 years, the U.S. National Assessment of Educational Progress (NAEP), like virtually all state assessments, has reported results mainly in terms of average performance. But NAEP has avoided statements about whether test performance was good enough or whether there were any facts or competencies that students at any particular grade should be expected to know. Now, many policy-makers want to change test score reporting - in the case of NAEP, they want to know the percent of students functioning at three performance levels: Basic, Proficient, and Advanced. For the last twelve months, measurement specialists have tried to implement the best available knowledge about standard setting to set national standards in mathematics (Hambleton & Bourque, 1991). The final standars were a qualified success, though a great deal was learned about national standard setting that has influenced a second study to set 1992 national standards in the areas of reading, mathematics, and writing.

Among the points that were learned from the first study are:

- The need for substantial technical consultation to address every aspect of the standard-setting process;
- The need for a random sampling design to select judges;
- Operationalized definitions of marginally basic, proficient, and advanced students; and
- Attention to intra-judge and inter-judge reliability of item ratings throughout the process, and procedures to aggregate data.

It remains to be seen whether the 1992 national standard-setting process will receive less criticism than the 1990 effort. In any case, measurement advances in national standard setting have increased the likelihood of success in 1992, and have resulted in the production of data that policy-makers have found useful.

Measurement Models

The development of measurement models for the design of tests and evaluation of test results is an on-going activity of psychometricians. The year of 1968 was especially important because of the publication of Lord and Novick's (1968) *Statistical Theories of Mental Test Scores*. This publication marked the beginning of the "modern era" of testing with the introduction of "item response theory".

Item response theory (IRT) was in no way inspired by policy-makers; however, it now seems clear that testing practices can be enhanced (i.e., test validity can be increased) with the successful applications of IRt in test development, identification of potentially biased items, equating of test scores, computerized adaptive testing (CAT), test score reporting, and diagnostic assessment (see, for example, Hambleton, Swaminathan, & Rogers, 1991).

Certainly, IRT is not a "magic wand" to wave over bad items, inadequate test designs, and poorly administered tests, but, in the hands of competent measurement specialists, positive enhancements to testing can result and thereby produce better data for educational policy-makers. Among the best examples are the IRT reporting scales with NAEP which provide policy-makers with highly informative outcomes of educational performance.

Translating Tests

Policy-makers currently have considerable interest in international comparative studies of educational progress. Many potential threats to test score validity exist in these studies, including choice of test content and item formats, and choice of examinee samples. One of the major technical problems to be overcome concerns the process of adapting tests for use in different cultures and languages and establishing score equivalence of the various forms of the tests. The upcoming IEA study of mathematics and science achievement will be carried out in over 60 countries. The scale of the test translation problem seems enormous but must be overcome if policy-makers are to have access to valid data.

Various technical problems are associated with the process of translating tests: (1) identifying and minimizing cultural differences, (2) identifying the appropriate language for testing, (3) finding equivalent words or phrases, and (4) locating qualified translators. Fortunately, each of these problems can be overcome in practice, but substantial time is required. Beyond test translation are the problems of choosing and implementing designs for establishing translation equivalence. Both judgmental and statistical designs exist though the former (e.g., backward translations) are more common. In view of the importance of establishing test score equivalence, however, the use of statistical designs and analyses seems essential. IRT seems to provide the most defensible designs because untenable assumptions about the equivalence of score distributions in the various language groups do not need to be made. Clearly, though, recent methodological advances have provided a technically sounder basis on which to translate tests and to establish test score equivalence (Hambleton & Bollwark, 1991).

Summary

Policy-makers today often want (and certainly need) test data to influence their decision making. As a result, many weaknesses in testing practices have been uncovered and highlighted (e.g., item and test bias; use of inappropriate item formats) to protect those who could be unfairly affected by invalid test results. Also, shortcomings in measurement knowledge with respect to the new demands of policy-makers (e.g., setting national standards) have also been uncovered. The result has been important advances in measurement technology and practice in areas such as item formats for testing, setting standards, measurement models, computerized testing, and translating tests. There is clear evidence that policy-makers and researchers are beginning to work together, but problems remain to be overcome. Researchers must learn to respect the practical constraints under which policy-makers work. Sometimes, too, researchers must be prepared to offer best judgments based upon incomplete data if they want to be influential in the decision-making process. On the other hand, policy-makers must be willing to fund basic research studies, and, at the same time, they must have realistic expectations about what researchers are capable of doing within a short time frame, and respect the fact that good psychometric work takes time and resources. With more understanding between researchers and policy-makers, there is every reason to expect that better educational decision-making will result.

References

Hambleton, R.K., & Bollwark, J.A. (1991). Adapting tests for use in different cultures: Technical issues and methods. Bulletin of the International Test Commission, 17, 3-32.

714

Hambleton, R.K., & Bourque, M.L. (1991). Initial performance standards for the 1990 NAEP Mathematics Assessment (Volume III, Technical Report). Washington, DC: National Assessment Governing Board.

Hambleton, R.K., & Murphy, E. (1992). A psychometric perspective on authentic measurement. Applied Measurement in Education, 5(1), 1-16.

Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). Fundamentals of item response theory. Newbury Park, CA: Sage.

Lord, F.M., & Novick, M.R. (1968). Statistical theories of mental test scores. Reading, MA: Addison-Wesley.

A MULTIPLE LINEAR REGRESSION APPROACH TO PLACEMENT AT THE UNITED STATES COAST GUARD ACADEMY

E. Manfred, & G. Rezendes, Mathematics Department, United States Coast Guard Academy, New London, USA

Abstract

Mathematics departments throughout the United States are concerned with the failure rate in freshman calculus courses. Studies by the Mathematics Association of America place the failure rate as high as fifty percent. Currently, there is considerable effort by mathematicians, educators and professional organizations to reverse this trend. It is clear that standardized test scores and high school grades are not a measure of calculus readiness. The Mathematics Department at the United States Coast Guard Academy has been using a multiple linear regression model to predict freshman calculus grades. The model uses all available data on incoming cadets as one part of a total placement procedure. Information gleaned from such procedures reduces errors stemming from underestimating or overestimating a student's capabilities and potential achievements in mathematics. Such errors may not only lead to failure, they preclude the best educational challenge.

The literature is replete with studies that predict term or cumulative grade point averages. Very little has been done for predicting grades in individual courses. The regression model includes standardized test scores, high school rank, number and type of mathematics courses taken in high school, an algebra achievement exam, a trigonometry achievement exam, plus several other independent variables, An R–Square of .48 indicates that approximately 48% of the variance in cadet grades is predictable from the model. The efficiency of the model will also be discussed.

716

AN ADAPTIVE TESTING MODEL FOR ASSESSING FACTORS IMPEDING SKILL ACQUISITION

George B Macready, & Ralph DeAyala, University of Maryland, USA

A measurement topic of much current research interest is *adaptive testing*, in which test items are selected and presented by means of a computerized system instead of in the conventional method of a preset list of items (see for example, Kingsbury & Weiss, 1983; Thissen & Mislevy, 1990). By selecting items near the performance or skill level of the examinee, major savings (e.g., 20% to 50%) in test length can be achieved without sacrificing precision of measurement. Many of the "mastery" decisions regarding diagnosis, placement, and/or screening are being made using computer adaptive testing. At present, the major theoretical foundation for adaptive testing is Item Response Theory (IRT), in which the underlying attribute being assessed is assumed to be a continuous trait (see for example, Lewis & Sheehan, in press; Wainer & Kiely, 1987; Weiss, 1982). However, when one is interested in assessing mastery of highly focused tasks or skills, the use of an IRT-based testing strategy provides an unrealistic modeling framework which carries with it a number of practical and theoretical disadvantages (see Macready & Dayton 1980). Recently, Macready & Dayton (in press) have extended the adaptive testing paradigm by suggesting the use of an alternative probability model which is based on a categorical latent variable. This alternative modeling framework may be particularly useful when an investigator is interested in assessing mastery of specific skills. Such diagnostic testing may be especially helpful in prescribing appropriate instructional remediation for individual examinees, while providing a substantial reduction in the number of items required for effective diagnosis. In a simulated adaptive testing application, Macready and Dayton were able to reduce the number of items required for use in assessment by about 50 percent when compared with a nonadaptive classification procedure. In addition, the classification outcomes they obtained were highly consistent with those obtained when all available items were used in making classification decisions.

However, the algorithm proposed by Macready and Dayton (in press) has a number of properties which potentially make it less efficient than what it might otherwise be. These include the following:

- it does not use all of the available i-tem response information in the adjustment of prior probabilities related to latent class membership,

- it requires that the sequence of item ,election be ordered by content domain (i.e., all items selected for administration from each domain must be considered consecutively, without the interjection of items from other domains) and
- it incorporates an empirically based procedure in the adjustment of priors which uses data from previously assessed examinees with equivalent classification decisions as those attained by the examinee being assessed. This may present a practical problem as the number of classification decisions becomes large since the number of equivalently classified examinees will in many cases rapidly diminish, resulting in potentially unacceptable inaccuracy in the estimated adjusted priors.

The purpose of this paper is to develop an alternative adaptive testing algorithm which is similar in purpose and form to that presented by Macready and Dayton (in press) but which circumvents some of the potential weaknesses we have identified with their approach. The procedure we propose is also based on a latent class modeling framework and might appropriately be considered for uses when one is interested in making classification decisions for each of K different latent categorical attributes. This proposed procedure incorporates an iterative process to determine which domains an examinee has acquired (i.e., mastered). The first step within each iteration is to decide which item from which domain will maximize the likelihood of proper classifications across the set of K domains of interest. Once the desired item is identified, it is administered and scored and the resulting outcome is used to establish "tentative" classification decisions for each of the K domains. These tentative classifications are deemed to be "final" if the estimated levels of misclassification are at or below a user specified maximum acceptable classification error rate. If the tentative classifications are not deemed to be final, than an additional item is selected based on the same criterion as above. This item is also administered and scored and once again the resulting classifications are assessed with respect to their acceptability in terms of error. This iterative process is continued until it is possible to make acceptable classification decisions for each domain. Note that, although the proposed algorithm provides for classification decisions on all of the domains of interest, items directly related to each domain need not necessarily be administered. This is particularly likely to be the case when there are strong relations among latent class membership across the domains of interest.

To assess the effectiveness of the proposed algorithm, a simulation based upon real data was conducted. Comparisons are made among the procedure here proposed, the procedure developed by Macready and Dayton (in press) and a nonadaptive procedure with respect to the expected number of items required for classification. In addition, the proposed algorithm is compared for consistency of classifications with outcomes obtained when all available items are used in making these decisions.

References

Macready, G.B. & Dayton, C.M. (1980). The nature and use of state mastery models. Applied Psychological Measurement, 4, 493-516.

Macready, G.B. & Dayton, C.M. (in press). The application of latent class models in adaptive testing. Psychometrika.

Kingsbury, G.G. & Weiss, D.J. (1983). A comparison of IRT-based adaptive mastery testing and a sequential mastery testing procedure. In D.J. Weiss (Ed.), New horizons in testing: Latent trait test theory and computerized adaptive testing (pp. 257-283). New York: Academic Press.

Lewis, C. & Sheehan, K. (in press). Using Bayesian decision theory to design a computerized mastery test. Applied Psychological Measurement.

Thissen, D. & Mislevy, R.J. (1990). Testing algorithms. In H. Wainer (Ed.), Computerized adaptive testing: A primer (pp. 65-102). Hillsdale, NJ: Erbaum.

Wainer, H. & Kiely, G.L. (1987). Item clusters and computerized adaptive testing: A case for testlets. Journal of Educational Measurement, 24, 185-201.

Weiss, D.J. (1982). Improving measurement quality and efficiency with adaptive testing. Applied Psychological Measurement, 6, 473-492.

719

SELF-ADAPTED TESTING: PROMISE AND IMPLICATIONS

Steven L. Wise, Department of Educational Psychology, University of Nebraska-Lincoln, USA

Introduction to the Problem

The purpose of this paper is to discuss an alternative method of computerized testing, termed self-adapted testing, for the estimation of examinee ability. Recent research has suggested that the use of a self-adapted testing format may reduce examinee test anxiety and lead to more valid measurement of student learning. In my paper, I will

introduce the theory underlying self-adapted testing,
review previous research relevant to this type of testing, and
discuss the potential benefits of self-adapted testing.

Theoretical Perspective

The development of Item Response Theory (IRT) has made it possible for the test performance of examinees to be compared on the same scale of measurement even if they are administered different sets of test items. According to IRT, examinee ability estimation is independent of the particular set of items administered from a calibrated pool. A popular application of this feature of IRT is a computerized adaptive test (CAT), in which a computer algorithm matches the difficulty of the items administered to the estimated ability of each examinee. At each step in a CAT, the next item to be administered is a function of the examinee's responses to previously administered items. Using a CAT, examinee ability is estimated more efficiently than with a conventional test because typically fewer items are required to attain the same degree of measurement precision. Although a primary advantage of a CAT is efficiency of ability estimation, claims are often made regarding the effects of a CAT on examinee motivation. When describing the benefits of adaptive testing, Hulin, Drasgow, and Parsons (1983) stated that, "motivational problems should also be reduced: low ability examinees will not be daunted by very difficult items and high ability examinees are not likely to become careless after answering several consecutive easy items." This claim, however, does not appear to have been investigated empirically.

Several years ago, Rocklin and O'Donnell (1987) explored an innovative application of IRT in computerized testing, termed *self-adapted* testing. In a self-adapted test, after each item is administered, the examinee is given feedback regarding the correctness of his/her answer and then is allowed to choose the difficulty of next item to be administered from among several (typically 6-8) levels

720

of difficulty. That is, the difficulty levels of the items administered are chosen by the examinee, rather than by a computer algorithm (as in a CAT). Rocklin and O'Donnell suggested that allowing examinees to choose the difficulty levels of their items should decrease feelings of anxiety that might hinder test performance.

Research Using Self-Adapted Tests

Rocklin and O'Donnell (1987) compared examinee performance on a self-adapted test with the performances of examinees taking two fixed-item tests from the same item pool, finding that the self-adapted test yielded a significantly higher mean ability estimate than the other tests. Wise, Plake, Johnson, and Roos (in press) compared the test performances of examinees who were randomly assigned to take either a self-adapted test or a CAT. They found that examinees taking the self-adapted test showed (a) significantly higher mean ability estimates and (b) significantly lower post-test state anxiety. Moreover, examinees taking the self-adapted test showed a strong tendency to choose item difficulty levels that were closely matched to their ability levels. Hence, self-adapted testing also appears to be reasonably effcient, though not quite as efficient as CAT. More recently, a study by Roos, Plake, and Wise (1992) investigated the importance of item feedback in self-adapted testing. They found that a self-adapted test yielded (a) significantly higher ability estimates than a CAT, even if item feedback is not given, and (b) significantly lower post-test state anxiety. The findings of both of these studies suggest that self-adapted tests may serve to lower examinee test anxiety, thereby reducing a debilitating influence on test scores and yielding higher ability estimates.

Implications for Educational Practice

According to the principles of IRT, a self-adapted test should not perform as well as a CAT in estimating examinee ability. Although both tests should provide unbiased estimates of an examinee's true ability, a CAT chooses items in a manner that is likely to be more efficient than that of an examinee, resulting in more precise ability estimation than with a self-adapted test. The results from the studies described above are intriguing because the higher mean ability estimates yielded by the self-adapted test are inconsistent with IRT principles. As a group, examinees in the self-adapted condition outperformed their randomly equivalent CAT counterparts, while choosing item difficulty levels similar to those that they would have received in the CAT condition. This implies that the items were generally easier for examinees in the self-adapted condition, an apparent violation of IRT parameter invariance. This inconsistency can be reconciled if one considers item parameters to be dependent on the context of test administration. That is, item parameters may be invariant across groups of examinees within a testing context, but not across contexts. If self-adapted testing reduces the overall anxiety level of a group of examinees, leading to increased test performance, then the difficulty parameters for at least some of the items should decrease. Note that this explanation implies that an examinee's success in passing an item is not simply a function of ability, but is also influenced by psychological factors, such as anxiety or motivation. The greater the influence of psychological factors on examinee test performance, the less complete are models based solely on ability (e.g., IRT).

If additional research confirms that a self-adapted test decreases test anxiety and consequently leads to higher test performance, then this testing format represents an innovative application of IRT that should yield more accurate estimates of examinee ability than a CAT. Moreover, a self-adapted test is more efficient than a conventional test; research has shown that examinees taking a self-adapted test tend to choose items that are very informative, from an IRT standpoint, for estimating ability. Hence, a self-adapted test may reduce a debilitating influence on test performance while providing efficient estimation of ability.

References

Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). Item Response Theory. Homewood, IL: Dow-Jones-Irwin.

Rocklin, T., & O'Donnell, A. M. (1987). Self-adapted testing: A performance-improving variant of computerized adaptive testing. Journal of Educational Psychology, 79(3), 315-319.

Roos, L. L., Plake, B. S., & Wise, S. L. (1992, April). The effects of feedback in computerized adaptive and self-adapted tests. Paper presented at the annual meeting of the National Council on Measurement in Education.

Wise, S. L., Plake, B. S., Johnson, P. L., & Roos, L. L. (in press). A comparison of self-adapted and computerized adaptive tests. Journal of Educational Measurement.

DESIGNING OF OPTIMAL SAMPLES FOR THE ESTIMATION OF PARAMETERS IN ITEM RESPONSE MODELS

Martijn P.F. Berger, Department of Education, University of Twente, Enschede, The Netherlands

Abstract

In this study criteria from optimal design research are applied to IRT models and the efficiency of IRT parameter estimates are compared for different sampling designs.

Two-stage and sequential sampling designs are compared with each other and some guidelines are offered for the designing of samples for the estimation of item parameters in IRT models.

Summary

The problem of designing experiments is encountered in various fields of research and several aspects have been studied in the statistical literature on optimal designs. Reviews of these developments are given by Atkinson (1982), Steinberg and Hunter (1984), and Ford, Titterington and Kitsos (1989). In the past few years the ideas from optimal design research have been applied to test theory and item response theory (IRT) models. See Berger and van der Linden (1991) for a review.

Designs are generally optimized by maximizing information on the parameters of an IRT model. Van der Linden (1988) used information about the ability parameters to optimally design tests. Lord and Wingersky (1985), Thissen and Wainer (1982) and de Gruijter (1985, 1988) used asymptotic variances of the estimators of the item parameters to compare relative efficiencies of tests and models. In so-called two-stage testing procedures efficiency is increased by matching abilities with item difficulties.

In situations where IRT models are applied to different groups, matrix sampling designs have been implemented. The results of Lord (1962) and Pandey and Carlson (1976) stress the importance of multiple matrix sampling designs to increase efficiency in estimating parameters. Berger (1991) and van der Linden (1989) investigated the efficiency of some sampling designs for IRT models and Vale (1986) minimized equating errors with sampling designs.

The purpose of the present study is to apply the criteria from optimal design research to IRT models and to compare the efficiency of the IRT parameter estimates for different sampling designs. Some results and guidelines will be given for the designing of samples for the efficient estimation of item parameters in logistic IRT models.

723

Optimal Design Criteria and IRT Models

To decide whether one design should be preferred over another, a criterion is needed. If an IRT model is considered for the test data, then a suitable criterion for the efficiency is based on some function of Fisher's information measure in the sample. Although several criteria have been proposed in the literature on optimal designs, the most familiar are the D-optimality, the E-optimality and the A-optimality criterion. These criteria can be applied to IRT models.

There is, however, a problem. Fisher's information matrix on the item parameters is generally not independent of the values of the parameters in the model. This means that we will have to know these values in advance, before we can select a design for the efficient estimation of these parameters. Several procedures to go around this problem have been proposed.

In this study three procedures are considered and compared with each other, namely, approximation of the parameters values, two-stage designs, and sequential designs. The results are limited to logistic IRT models and based on different values for the item parameters.

Results and Conclusions

Approximation of parameter values

Since the effective range of the item parameters is usually limited, it may be possible to get a good overall picture of design efficiency by computing efficiency for combinations of parameter values. Such an approach was followed by Berger (1991). The results of this study support his conclusions and indicate that the optimality criteria are quite robust to poor initial estimates.

Two-stage designs

The second procedure to avoid the problem of unknown parameter values is to divide the testing process into stages. In this form of testing, initial estimates of item parameters are updated at each stage. The results of this study show that two-stage testing greatly improves the efficiency of the parameters, especially when the initial estimates are poor.

Sequential designs

The last procedure considered in this study is a sequential design procedure. This procedure not only generates an optimal sampling design sequentially, but also makes it possible to estimate the IRT parameters sequentially. The results show that this procedure is consistent and very efficient in finding optimal designs.

References

Aitkinson, A.C. (1982). Developments in the design of experiments. International Statistical Review, 50, 161-177.

Berger, M.P.F. & van der Linden, W.J. (1991). Optimality of sampling designs in item response theory models. In: M. Wilson (ed.) Objective Measurement: Theory into Practice. Norwood: Ablex Publishing Company.

Berger, M.P.F. (1991). On the efficiency of IRT models when applied to different sampling designs. Applied Psychological Measurement, 15, 293-306.

Berger, M.P.F. (In press). Sequential sampling designs for the two-parameter item response theory model. Psychometrika.

Ford, I., Titterington, D.M. & Kitsos, C.P. (1989). Recent advances in non linear experimental design. Technometrics, 31, 49-60.
de Gruijter, D.N.M. (1985). A note on the asymptotic variance-covariance matrix of item parameter estimates in the Rasch model. Psychometrika, 50, 247-249.
de Gruijter, D.N.M. (1988). Standard errors of item parameter estimates in incomplete designs. Applied Psychological Measurement, 12, 109-116.
Lord, F.M. & Wingersky, M.S. (1985). Sampling variances and covariances of parameter estimates in item response theory. In D.J. Weiss (ed.) Proceedings of the 1982 Item Response Theory and Computerized Adaptive Testing Conference, pp. 69-88, Minneapolis: University of Minnesota.
Lord, F.M. (1962). Estimating norms by item-sampling. Educational and Psychological Measurement, 22, 259-267.
Steinberg, D.M. & Hunter, W.G. (1984). Experimental design: Review and comment. Technometrics, 26, 71-130.
Theunissen, T.J.J.M. (1985). Binary programming and test desing. Psychometrika, 50, 411-420.
Thissen, D. & Wainer, H. (1982). Some standard errors in item response theory. Psychometrika, 47, 397-412.
Vale, C.D. (1986). Linking item parameters onto a common scale. Applied Psychological Measurement, 10, 333-344.
van der Linden, W.J. (1988). Optimizing incomplete sampling designs for item response model parameters (Research Report 88-5). Enschede: University of Twente.
van der Linden, W.J. (1989). Optimizing item calibration designs. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.

OPTIMAL SAMPLING DESIGNS FOR GROUP-LEVEL IRT MODELS

Hilde Tobi, Department of Education, University of Twente, Enschede, The Netherlands

Introduction

Assessment studies have increased interest in models for analyzing grouped data. This has resulted in the formulation of Item Response Theory (IRT) models which can be used for grouped data (Bock & Mislevy, 1981; Mislevy, 1983; Reiser, 1983). In many applications of IRT models the use of sampling designs can increase efficiency. This paper handles about the efficiency of designs for one class of group-level IRT models, the group-fixed effects models.

Group-level IRT

A group-level IRT model describes the probability of obtaining a correct response by an at random selected member of a specified group. By groups are meant salient subpopulations which can be identified by such variables as sex, race and urbanity. The group-level model used in this study is a model that was first introduced by Reiser (1983), but has been rewritten to fit comparison with the individual-level two-parameter logistic IRT model. The group-fixed effects model considers variability at the examinee level as independent error within the lowest subclass of the classification (the group). So the model incorporates just one ability parameter for each group.

Group-level IRT models can be used for the assessment of groups of examinees in large scale testing programs. In this type of programs efficiency is very important, because of the usual lack of classroom time and money. Matrix sampling is a technique to help optimize efficiency.

Efficiency of sampling designs

The central idea of matrix sampling is very straightforward: it is not necessary to administer every item from an item domain to each examinee of a group if one only wants to estimate the performance of that group on the item domain.

Berger and van der Linden (1991) give a review of various ways to summarize and compare the efficiency over different designs. In this study one of these criteria is used: Fisher's information (Lindgren, 1976, p.249). The trace of Fisher's information matrix is taken as the measure of efficiency of a design. The literature (Gressard & Loyd, 1991; van der Linden, 1988; Berger, 1989, 1991) indicates that results on efficiency of designs depend on the range of the item parameters.

An other type of efficiency is cost efficiency. A less expensive design which yields an equal amount of information is more cost efficient. When the amount of information is about equal for different designs, a cost criterion can help to select the optimal design.

Method

Three tests are generated which differ with regards to the range of the item slope and item difficulty parameter. These tests can be regarded as a mastery test, general achievement test and an admission test. For each of these three tests the test length is fixed at 40 items. For each of the three tests four designs were considered: the complete design for the individual-level model and three sampling designs for the group-level model. These three designs have different patterns for the examinee-item combinations but the total number of observations will be kept constant over the test-design combinations.
It is assumed that the population of examinees consists of five different groups with ability levels -2, -1, 0, 1 and 2, respectively. Within each group the abilities are normally distributed with unit variance. The item parameters as well as the ability parameters of the examinees for the individual-level IRT model were generated by routines from a mathematical library, and -when necessary- rescaled.

Results

The results of the simulation for the three types of tests did not show one uniformly the best design with regard to information on the group ability parameter. Sometimes one design gives more information than the others on one specific value of the group-ability parameter, but less information on the others.
With regard to information on item parameters, it is also hard to tell which design is the best, even if information on one of the item parameters is considered more important than information on the other.
However, with regard to costs the four designs differ significantly. During the paper presentation it will be shown that differential costs of the designs can be parametrized. Dependant on the relative importance of specific cost factors, the less expensive design can be chosen.

Discussion

For the tests generated and the designs used in this study, keeping the total number of observations constant, the distribution of parameters seem to be of more importance to the obtained information than sampling design.
A cost criterion can help to choose among designs which result in about the same quantity of information. A well defined cost function enables the researcher to choose the design which yields the most information for a specific amount of money.

References

Berger, M.P.F. (1989). On the efficiency of IRT models when applied to different sampling designs (research Report 89-4). Enschede: University of Twente.
Berger, M.P.F. (1991). On the efficiency of IRT models applied to sampling designs. Applied Psychological Measurement, 15, 293-306.

Berger, M.P.F. & Van der Linden,W.J.(1991). Optimality of sampling designs in item response theory models. In: M. Wilson (ed.) Objective Measurement: Theory into Practice. Norwood NJ: Ablex Publishing Corporation.

Bock, R.D. & Mislevy,R. (1981). An item response curve model for matrix sampling data: the California Grade-Three Assessment. New Directions for Testing and Measurement, 10, 65-89.

Gressard, R.P. & Loyd,B.H. (1991). A comparison of item sampling plans in the application of multiple matrix sampling. Journal of Educational Measurement, 28, 119-130.

Hambleton, R.K. & Swaminathan, H. (1985). Item response theory: principles and applications. Boston/Dordrecht/Lancaster: Kluwer-Nijhoff Publishing.

Lindgren,B.W. (1976). Statistical Theory. New York: MacMillan Publishing Co.

Mislevy, R.J. (1983). Item response models for grouped data. Journal of Educational Statistics, 8, 271-288.

Reiser, M. (1983). An item response model for the estimation of demografic effects. Journal of Educational Statistics, 8, 165-186.

Van der Linden, W.J. (1988). Optimizing incomplete sample designs for item response model parameters (Research Report 88-5). Enschede: University of Twente.

ADAPTING INFEASIBLE TEST CONSTRUCTION MODELS

Ellen Boekkooi-Timminga, University of Groningen; & Jos J. Adema, PTT-Research, Leidschendam, The Netherlands

Introduction

Recently, several linear programming methods for selecting optimum tests from item banks have been developed (e.g. Adema, Boekkooi-Timminga & van der Linden, 1991; Theunissen, 1985; van der Linden & Boekkooi-Timminga, 1989). One of the problems not taken into account, is the problem of not being able to find a test fitting all test requirements. In this case no test is selected. The test constructor has to guess which requirements caused the problem, and re-start the test construction process. Because, it may be expected that a test constructor tries to formulate non-contradictory constraints that are expected to be realizable, it is clear that adapting the test specification may be a very difficult task.
In this paper an overview of possible ways to adapt a test specification inheriting infeasibility is given. A procedure is proposed that checks whether a solution fitting all requirements can be found. If not, it guides the test constructor in adapting his/her test specification such that a feasible solution can be obtained. The procedure can be applied to all test construction models proposed in the literature.

Adapting Test Specifications

Mathematical programming models for test construction consist of an objective function, and a set of constraints that has to be fulfilled. The decision variables x_i used in the models are usually of the 0-1 type, and denote whether an item is ($x_i = 1$) or is not ($x_i = 0$) selected for the test. The following test specification is used for illustration:

(1) Maximize y

subject to

$$(2)\ \sum_{i=1}^{I} I_i(\theta_k)x_i - y \geq 0, \qquad k = 1, \ldots, K,$$

$$(3)\ \sum_{i=1}^{I} x_i = N,$$

$$(4)\ \sum_{i \in V_1} x_i = N_1,$$

$$(5)\ \sum_{i \in V_2} x_i = N_2,$$

$$(6)\ b_i x_i \leq \text{maxb} \qquad i = 1, \ldots, I,$$

$$(7)\ x_i \in \{0, 1\}, \qquad i = 1, \ldots, I.$$

Objective function (1) and the set of constraints in (2) state that the minimum amount of test information obtained at the ability levels k should be maximal. Constraint (3) requires that the test consists of N items. Furthermore, the test should contain N_1 items from subset V_1, and N_2 items from subset V_2 (Constraints [4] and [5]). The set of constraints in (6) requires that the item difficulty values b_i of the items in the test should be less than maxb.

There are several ways to adapt test specifications like (1) - (7), if they are infeasible. The basic cause of infeasibility is that the set of test requirements is in contradiction with the characteristics of the item bank. For example, the test specification states that the test should consist of 20 items, and that the test information has to be at least 10 at a certain ability level, while the maximum amount of test information that can be obtained at this ability level for 20 items is 9. In case there is no feasible solution the feasible region, i.e. the set of feasible solutions, is empty. The purpose of adapting a test specification is to create a feasible region. Boekkooi-Timminga and Adema (1992) give a detailed discussion of the possibilities. Some possible ways to adapt a test specification consider the following:

- Right-hand-side coefficients. In the above model the right-hand-side coefficients are, respectively: 0 for all constraints in (2), N in (3), N_1 in (4), N_2 in (5), and maxb for all constraints in (6). Changing the right-hand-sides in (2) will not increase the feasible region, because these constraints do not restrict this region in any way. An adaptation of N, N_1, and/or N_2 might help, however, it is not obvious how to change them (increase or decrease). On the other hand, increasing maxb may create a feasible region, if the infeasibility is due to this constraint.
- Equalities vs inequalities. Another option is to replace equalities by inequalities. Replacing inequality by equality signs is not recommended, because this may reduce the size of the feasible region. Carefulness is recommended if equality signs are used, for example, there is an extremely large chance that no solution can be found, if it is required to construct a test with a mean item difficulty of exactly 0.5. So equality signs should only be used in constraints that regard integer valued variables and constants.

- Item sets. It is also possible to change the set of items, like V_1 or V_2, considered in a certain constraint. Only an increase of the number of items in a set may enlarge the feasible region, a decrease may reduce the size of the feasible region. Whether it actually enlarges or reduces the size of the feasible region, and to what extend, depends on the other constraints in the test construction model.
- Delete constraints. It is possible that deleting a constraint causes an increase of the feasible region.
- Objective function vs constraints. If the same item characteristics are considered in the objective function and in one or more constraints, it is adviced to reconsider this carefully. The constraints might cause that the objective function is not active, and thus redundant, or that the constraints are not active, and thus redundant. In this case the test constructor has to decide whether the constraints should be deleted or a new objective function has to be formulated.

Method

In short the method is as follows: It is first checked whether there is a feasible solution to the relaxed 0-1 LP test construction model. The relaxed model is obtained by relaxing the constraints on the decision variables x_i into: $0 \leq x_i \leq 1$. If there is no solution to the relaxed problem, there is also no solution to the 0-1 model. Solving the relaxed problem, information can be obtained identifying the constraints that are most likely to have caused the problem of not being able to find a solution. Details on this proces can be found in Boekkooi-Timminga & Adema (1992). In fact the obtained information indicates for which constraints the right-hand-sides have to be changed, such that there is a larger change of finding a solution. Actually, the problematic constraints are ordered with respect to their relative importance of not being able to find a solution. This information is presented to the test constructor, who next can adapt the test specification.

Results

The method works well. It turns out that for most practical applications only a few constraints are indicated as problematic. Adapting the constraints that are indicated as *most* problematic is sufficient in most cases.

References

Adema, J.J., Boekkooi-Timminga, E., & van der Linden, W.J. (1991). Achievement test construction using 0-1 linear programming. European Journal of Operational Research, 55, 103-111.

Boekkooi-Timminga, E., & Adema, J.J. (1992). An interactive approach to adapting infeasible linear programming models for test construction. Groningen: University of Groningen.

Theunissen, T.J.J.M (1985). Binary programming and test design. Psychometrika, 50, 411-420.

van der Linden, W.J. & Boekkooi-Timminga, E. (1989). A maximin model for test design with practical constraints. Psychometrika, 54, 237-247.

MODELLING THE CONNECTION BETWEEN INDIVIDUAL BEHAVIOR AND MACRO-LEVEL OUTPUTS: UNDERSTANDING GRADE RETENTION, DROP-OUT AND STUDY-DELAY AS SYSTEM RIGIDITIES

Ben Wilbrink, Center for Educational Research, University of Amsterdam, Amsterdam, The Netherlands

Certain unwanted states of affairs in education are extremely resistant to change, as certainly is the case for grade retention in secondary education and for attrition and study-delay in higher education. The individual student might prevent flunking of grades or failing the examination by stepping up his or her investment of time and attention. The catch, of course, is that what is possible at the individual level, is not possible at the group- or system level. This is what history teaches us, whether we understand it or not. Doornbos (1985) points out that teachers in their behavior are restrained by the characteristics of the system they are working in: the problem of retention is a problem at the level of the educational system as system.
Evaluation of education should reflect this systemic character of education, and of most or all of its characteristics that currently are proposed as performance indicators. The problem then is that there are no known methods of data analysis or educational measurement that are fit to the task to elucidate this kind of systemic phenomena. Worse, because there are no adequate evaluaton instruments, it is not possible for the educational researcher to give policy makers a handle on these kinds of problem.
The general problem in the educational field is how individual behaviours connect to macro-educational characteristics, with the direction of influence obviously going both ways. Recently Coleman (1990) presented a theory of social systems that connects behavior of actors (students and teachers, for example) at the micro-level with phenomena occurring at the macro-level of the social system involved. The theory of Coleman has its roots in micro-economics, and is conceptually very different from traditional methodological approaches to social (and educational) phenomena. The paper explores the possibility to apply this theory to the mentioned problems in education.

Coleman's Social System Theory

The basic idea is that actors within a social system exchange goods in ways that maximize satisfaction with the total number of goods the actor controls (possesses) after exchange. Coleman uses the basic premise of micro-economic theory about human behavior: "the more the person has of a good, while remaining at the same

level of satisfaction (because of having less of something else), the less of any other good he will be giving up to get still more of the good." Coleman uses these simple but strong assumptions from micro-economic theory to model the transactions taking place in the social system considered, and to present a mathematical structure isomorphic to the structure of transactions. Given some variables (e.g. marks obtained, time spent) on which empirical data are available, the model allows estimation of a conceptually very different set of variables, some of which are characteristics of individual actors (e.g. their interest in marks obtained and in alternative uses of their time), others are characteristics of the system itself (e.g. the exchange rate of marks and time). The model is a system model because exchange of resources takes place within a specified group of actors, and every exchange is dependent on the other exchanges taking place in this system. Individual exchanges within the system are competitive. Now 'competitive' is a label that excellently fits most educational systems, as many educational philosophers (for example Erasmus) and researchers have remarked. Here we have a model that explicitly connects the micro-level with the macro-level in a meaningful way, i.e. not simply by accumulating individual data as happens in linear models analyses (correlation and regression, LISREL, multi-level).

Illustrative application to higher education data

Educational systems can be modelled by Coleman's theory. As an example data for the first year examination in the department of Law at the University of Amsterdam will be analysed. In a companion paper (in the Higher Education theme) this application and its implications for educational policy are reported on more fully. The resources being exchanged here are marks (from teachers to students) and time spent (from students to teachers). At the system level it is evident from the raw data that students are investing in their study only half of the 1600 hours they are expected to (by law), and only a minority have passed the first year examination 'in time'. The situation is stable, in the sense that over the (observed) years no appreciable changes in the statistics have taken place. The amazing thing is that the individual student is almost always able by a relatively small extra investment to pass the examination in time, should he or she wish to do so. However, the number of students wishing to do so must be small, witness the data. (The paper will present the basic statistics here.) Available data allow determination of the exchange rate of marks against time, a macro-level characteristic of this educational system, and of the interest of every student (and teacher) in marks and in time (to be spent in study or alternatively outside). This kind of analysis shows why everybody in this system acts as he or she acts, given the characteristics of the system, i.e. in this case given the exchange rate of marks and time. There are many details that might be discussed, but the main point is that this kind of analysis shows every actor within the system acting under influence of important system characteristics. Traditional techniques of data analysis (experimental and correlational methods) treat every individual observation as independent of other observations, denying from the very start the possibility of dependent, competitive, 'systemic' behaviours.

Discussion of implications and possibilities

Every policy maker interested in improvement of attrition, in higher motivation of students, or in better marks, must have insight into the ways in which the educational system influences individual behaviors en in turn is shaped by collective individual behaviour. Especially government policies are directed at effects at the system level. Failure to evaluate the effects of education as obtained under the particular constraints of the system as (traditionally) implemented will lead to failure of law-induced educational innovation.

Social system analysis using marks and time expenditure will expose the weaknesses of current methods of educational assessment. The current marking system is taken for granted by almost every author on the subject. De Groot (1966) severely criticized the Dutch marking system, but he failed to indicate the historical roots of this particular marking system, using a scale from 1 to 10, all marks below 6 being unsatisfactory. Also in American handbooks of educational measurement no mention is made of the historical underpinnings of the current marking system in the U.S.A.. Chapman (1988), however, does present a historical survey of the American grading system, showing intelligence testing to be an offspring of the movement of standardized testing in education.

A social system where marks are exchanged against time is extremely vulnerable to the weaknesses of assessment and grading. The scale of marks is extremely coarse, not allowing any fine-tuning between amount of study time and mark received. Educational testing itself functions as a kind of 'clearing house', to make possible the exchange of time and marks, marks being the valued scores of the tests used. Regrettably educational measurement experts, with few exceptions, as for example Van Naerssen (1974) on models of strategic behavior of students, are not concerned with this 'clearing house' function of tests. The whole of item response theory is irrelevant to this 'clearing house' function of educational testing. It is quite possible that poor predictability of marks (not necessarily the same as poor predictability of test scores; predictability by the student is meant here) is the main reason for the observed exchange rate for marks against time in the law education case. Certainly the social system analysis in this case indicates there is plenty of room for improvement using modeling techniques for educational assessment as developed by Van Naerssen (1974) and Wilbrink (1978). Also in secondary education a social system analysis approach could point out the classical marking system as the main culprit in the case against (Dutch) retention rates of 15 to 25 percent every year. Remarkably, in English speaking countries there is an 'educational reform' movement in the direction of stricter grading policies (Shepard & Smith, 1989), in the direction also of more sorrow for more students.

Social system analysis re-establishes a fact that seems to be known to everyone except specialists in educational assessment: the difference between a psychological test and an educational test is that for the first the student prepares by sleeping well, and for the second by studying until a personally satisfactory level of mastery is reached. The characteristics of the educational system determine what the student thinks is a satisfactory level. Change this characteristic of the system, and the effectiveness of schooling will change.

References

Chapman, P. D. (1988). Schools as sorters. Lewis M. Terman, applied psychology, and the intelligence testing movement, 1890-1930. New York: New York University Press.

Coleman, J. (1990). Foundations of social theory. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Doonbos, K. (1985). Naar een macro-educatieve analyse van het zittenblijven en verwante expulsiefenomenen in het Nederlandse schoolwezen. In Wald, A., Een jaartje overdoen. Verslag van het SVO-symposium over zittenblijven in het voortgezet onderwijs(pp. 35-67). Lisse: Swets & Zeitlinger.

Naerssen, R. F. van (1974). A mathematical model for the optimal use of criterion referenced tests. Nederlands Tijdschrift voor de Psychologie, 29, 431-445.

Shepard, L. A., & Smith, M. L. (Editors) (1989). Flunking grades. Research and policies on retention. London: The Falmer Press.

Wilbrink, B (1978). Studiestrategieën. Amsterdam: Stichting Centrum voor Onderwijsonderzoek.

DETERMINING THE OPTIMAL MEASUREMENT DIRECTION IN A MULTIDIMENSIONAL LATENT SPACE

Terry A. Ackerman, University of Illinois, Illinois, USA

Abstract

Most standardized tests contain items that are capable of measuring different composites of multiple skills. This can become a problem if the measurement of multiple skills is reported as a single score. For a multidimensional test the reporting of a single score, and thus the explicit ordering of examinees, implies that the multiple skills being assessed are being represented as a single composite skill described by a unidimensional continuum or direction in the multidimensional latent ability space. The purpose of this paper is to examine two approaches for determining this direction, termed the "optimal measurement direction", for a multidimensional test. How knowledge of these directions can be applied to practical testing problems such as equating and item bias is discussed.

In most testing situations practitioners assume that, based upon the items written to match a given table of specifications, that they can identify the composite of skills being measured by a test. If a single score is to be reported it is, it is further assumed that all of the items in the test are either measuring the same skill or the same composite of skills. Realistically this is too strong of an assumption. However, in reality it is the interaction of the multidimensional distribution of cognitive skills of the examinees and the measurement capability of the items which ultimately determine the dimensionality of the complete latent space. Likewise, it is also this interaction that determines which composite of skills or direction in the complete latent space will be represented by reporting only one score for each examinee. This direction is referred to as the *optimal measurement direction.*

If two groups of examinees have different underlying multidimensional ability distributions and the items are capable of distinguishing between levels of these abilities, the unidimensional score scale could represent a different composite of skills for each group. That is, the interpretation of the skills being ordered by reporting only a single score for all examinees could be substantively quite different.

It is the purpose of this paper to examine two different methods to compute the direction assumed by a single score scale for a multidimensional test. After presenting the two procedures, applications of the optimal measurement direction to test equating and item bias detection will be illustrated using several examples.

Wang (1986) determined an explicit algebraic relationship between the true multidimensional ability distribution and the true multidimensional item parameters and the (assumed) unidimensional 2PL item response theory (IRT) model. She demonstrated that when a unidimensional IRT estimation program using maximum likelihood estimation procedures to calibrate multidimensional data, the resulting estimated latent ability scale, termed the *reference composite, is* actually a weighted composite of the underlying multiple dimensions. Specifically, the reference composite is essentially a weighted standardized composite of the underlying multiple abilities for which the test attains the maximum expected

Fisher information, $I\left(\hat{\theta}|\theta\right)$. Wang described the reference composite as "the primary ability that the test is intended to be measured. "

Ackerman (1991) reporting on collaborative research with Dr. William Stout and Dr. Brian Junker presented another approach. This approach assesses the ordinal accuracy of reporting a single score in a specified direction in a two-dimensional latent space. They determined how well the observed score X correctly orders examinees whose true order is given by a number correct score, γ, on some composite of the multiple skills (θ_1,θ_2) being assessed by the test. The approach suggested by Ackerman et al, was to determine which composite direction would make

$$P[X^{(2)} > X^{(1)} \mid \gamma^{(1)},\gamma^{(2)}]$$

as close to 1 as possible for all $\gamma^{(1)}$, $\gamma^{(2)}$ examinee pairs for which $\gamma^{(2)}>\gamma^{(1)}$. Note that $\gamma^{(i)}$ and $X^{(i)}$ represent the latent composite and the observed score, respectively, for examinee i. By examining the average value of $I(X|\gamma)$ over a given range of γ for various composites (i.e., directions corresponding to increasing angles from 0 to 90 degrees) the direction of maximum number correct information can be determined.

Preliminary findings from comparison studies suggest that the approaches yield very similar results. Interestingly when data is simulated for two groups and yields quite discrepant measurement directions, analyses which rely on the reliability and interpretation of the observed score scale, (i.e., Mantel Haenszel item bias analyses or equipercentile equating) can be greatly effected.

Ackerman (in press) pointed out that bias analyses which condition on observed scores which represent two different measurement directions are in reality comparing "apples to oranges". That is, bias analyses like the Mantel Haenszel which require

2 X 2 contingency tables for each possible score category provide a comparison of performance for the same *numerical* categories but potentially quite different *substanhve* categories.

Preliminary findings further suggest that the measurement direction for a particular test can change when equipercentile equating is applied. That is, the measurement direction implied by the number correct scale can be different from the direction suggested by the scaled scores.

References

Ackerman, Terry A. (1991, November). Measurement direction in a multidimensional latent space and the role it plays in bias detection. Paper presented at the 1991 International Symposium on Modern Theories in Measurement: Problems and Issues. Montebello, Quebec.

Ackerman, Terry A. (in press). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. Journal of Educational Measurement.

Wang, Ming M. (1986, April). Fitting a unidimensional model to multidimensional item response data. Paper presented at the Office of Naval Research Contractor's Meeting, Gatlinburg, TN.

USING RESTRICTED FACTOR ANALYSIS WITH BINARY DATA FOR ITEM BIAS DETECTION AND ITEM ANALYSIS

Frans J. Oort, University of Amsterdam, Amsterdam, The Netherlands, & Giray Berberoglu, Middle East Technical University, Ankara, Turkey

In Psychology (and education) there are all sorts of tests measuring abilities, aptitudes, personality traits, attitudes, moods, emotions, and so on. To be valid, each test scale should be unidimensional and free of item bias. Using the common factor model as an item response model, restrictive factor analysis (RFA) can be used to detect item bias, to test for unidimensionality, and to replace ordinary item analysis. Of course factor analytic procedures are really only suited to be used with items with continuous or interval response scales. Still, there are ways round the problem of factor analyzing non–continuous data, so that RFA can also be used with items with ordered categorical or dichotomous response scales.
In the present paper RFA is used to detect bias in dichotomous items and to carry out an item analysis. Data are from 1293 Turkish students who completed a 21 item intelligence test. Bias is investigated with respect to student's sex and school grade. Subsequently, item analysis is used to further shorten the intelligence test to become unidimensional and efficient. Finally, the RFA results will be compared to the results of Berberoglu & Oort (1992), who analyzed the same data using Birnbaum's logistic two–parameter item response model.

Method

Subjects are 1293 Turkish students from Ankara. 732 students are in the lower education level (junior high school, grades 4 to 6) and 561 students are in the higher education level (high school, grades 7 to 9). There are 686 males and 507 females.
The intelligence test to be analyzed is a Turkish version of the Group Assessment of Logical Thinking (GALT) test (Aksu, Berberoglu, and Paykoç, 1991). The GALT test consists of 21 items measuring six different logical skills. All items are scored 0 (false) or 1 (correct).
Psychometric analyses are done within the framework of the Theory of Violators, that is, the GALT test will be shortened to become unidimensional and efficient through item bias detection and item analysis (Oort, 1991a, 1991b). All analyses are done by means of RFA. As the items are dichotomous, the matrix of tetrachoric correlations is analyzed with the weighted least squares method (Jöreskog & Sörbom, 1988).

733

Item bias detection is done in the way described by Oort (1991a, in press), that is, items and potential violators of item purity (i.e., sex and school grade) all are accommodated in a factor model. Modification indices (MI) and expected parameter changes (EPC) serve as indicators of item bias. Subsequently, RFA is applied once more to do a general item analysis, that is, investigate local independence and calculate the item information index (described by Mellenbergh, 1991).
To compare the RFA results with the results from using Birnbaum's logistic two–parameter item response model (reported by Berberoglu & Oort, 1992), rank correlations are computed for item parameters and bias indices.

Results

According to the AMI decision rule (Oort, in press), only one of the items appears to be (moderately) biased with respect to sex (against females) and none of the items are biased with respect to school grade. However, there appear to be five pairs of items that violate local independence, as is indicated by high MIs and EPCs for residual covariances.
After removing the supposedly biased item and one of each pair of items that violates local independence, 15 items remain. In order to further shorten the test, information indices are calculated. Three items have very low information values and are removed from the test. To make it a ten item test, two items with moderate information values are removed because they measure a logical skill that is over–represented in the shortened GALT test.
Berberoglu and Oort (1992) used the perhaps more appropriate Birnbaum model on the same data and also constructed a ten item short form of the GALT test. Comparing both ten item tests it appears that they have eight items in common. Still there are differences. With Birnbaum, six items are rejected because of bias (RFA: one item) and none because of local dependencies (RFA: five items). Maybe Berberoglu and Oort applied too strict criteria for item bias. More noteworthy is that it looks like the Birnbaum model is insensitive to violations of local independence.
As a measure of the extent to which the RFA and Birnbaum models agree, rank correlations are calculated between corresponding item parameters, information indices, and item bias indices. The difficulty index, discrimination index, and the maximum information values that result from fitting the Birnbaum model agree largely with the corresponding indices in RFA (i.e. mean, factor loading, information index). The correlations are .99, .99, and .97 respectively (the number of observations is 21, i.e. the number of items). As for the item bias indices, the highest rank correlations are found between the EPC of RFA and the weighted signed area index: .94 with sex bias and .92 with grade bias.

Discussion

This exercise shows that RFA is very suitable to analyze dichotomous items. The RFA results agree to a very large extent with the results of fitting an item response model that seems more appropriate for the analysis of dichotomous items. Moreover, using RFA it is possible to detect violations of local independence. Other advantages of using RFA are that it is possible to detect bias with respect to

several potential violators simultaneously and that there is no need to group subjects into subgroups.
However, RFA has its disadvantages too. Firstly, only linear relations can be examined, making it impossible to detect nonuniform bias and nonlinear violations of local independence. Secondly, RFA will only detect item bias when the bias is present in the sample at hand. That is why RFA item bias indices show the highest agreement with the weighted signed area index and only moderate agreement with the unweighted unsigned area index.

References

O', Berberoglu, G. & Paykoç, F. (1991). Can the GALT test be used in a different cultural setting? Research Report Middle East Technical University, Ankara, Turkey.

Berberoglu, G. & Oort, F.J. (1992, June). Psychometric Analysis of the Turkish Group Assessment of Logical Thinking (GALT) Test. Paper presented at the European Conference on Educational Research, Enschede, The Netherlands.

Jöreskog, K.G. & Sörbom, D. (1986). *LISREL 7.* A guide to the program and applications. Mooresville, IN: Scientific Software, Inc.

Mellenbergh, G.J. (1991, July). The unidimensional latent trait model for continuous item responses. Paper presented at the 7th European Meeting of the Psychometric Society, Trier, Germany.

Oort, F.J. (1991a). Constructie van een eendimensionale schaal volgens de Schenderstheorie. [Using restricted factor analysis to construct a unidimensional scale in accordance with the Theory of Violators]. In: Hoogstraten, J., & Linden, W.J. van der (Eds). ORD' 91: Methodologie. Amsterdam: Stichting Centrum voor Onderwijsonderzoek van de Universiteit van Amsterdam.

Oort, F.J. (1991b, July). Theory of Violators: Assessing Unidimensionality of Psychological Measures. Paper presented at the 7th European Meeting of the Psychometric Society, Trier, Germany.

Oort, F.J. (in press). Using restricted factor analysis to detect item bias. Methodika.

PHYSICAL DIMENSIONAL THEORY AS A METHOD FOR ANALYSIS OF PSYCHOLOGICAL DATA

Svetoslav Marinov, Institute of Psychology, Russian Academy of Science, Moscow, CIS

In mathematical analysis of psychological phenomena some specific problems have been existing which are typical for this research area. Problems of this type are, for example, a large number of variables which influence on the observed phenomenon, an absence of any analytical relations among these variables, a vague nature of these variables (absence of common scales, measurement units etc.). The existence of these problems is an explanation of the wide use of statistical methods as correlation and factor analysis in mathematical psychology. In our study we suggest some new ways for application of Physical Dimensional Theoty (PDTh) in analysis of psychological phenomena. At the present time this theory has wide application in many branches of the physics and engineering, but unfortunately, it is practically unknown for the researchers involved in analysis of psychological data. PDTh is based on the fundamental property of the physical laws (respectively of the physical phenomena) to be invariant relative to the choice of measurement units and physical scales. This theory gives good results in analysis of phenomena which depend on a large number of variables, and where are no analytical connections among these variables, i.e. in situations that occur frequently in analysis of psychological phenomena.

We see the following possibilities for application of the PDTh in analysis of psychological data. A traditional use of this theory permits:

- to reduce significantly the number of measurements and/or calculations necessary for analysis of certain phenomenon under examination;
- to make a conclusion regarding whether all significant variables are included in the description of the phenomenon or not;
- to determine some specific states, for which a fixed number of variables describes the phenomenon quantitatively, i.e. to determine the boundaries of qualitative homogeneity. An use of some inverted statements of PDTh permits;
- to determine in physical units the dimensions of variables which have a priori indefinite dimensions. This last possibility gives, we believe it, a new technics for analysis of psychological phenomena.

In our study some examples of psychological data analysis are reported as an illustration of the PDTh methods. Original and published earlier experimental data are involved. Three groups of data are analysed: standart deviation (SD) as a function of mean reaction time (RT) under different stimulus conditions; visual RT as a function of stimulus size (S) and distance (D) from a central fixation point in

the visual field; and proportion recalled (PR) as a function of retention interval (RI) and number of presentations (NP) at a fixed time interval in an experiment of Brown-Peterson type. All illustrations demonstrate a high efficiency of data processing and analysis by the PDTh methods. It is obtained a good quantitative agreement between our results and the results reported in other publications regarding the same data, but achieved by absolutely different methods.
It is concluded that the PDTh can be applied adequately and efficiently to the analysis of psychological phenomena.

TWENTY–FIVE YEARS IN THE TESTING TRENCHES: LESSONS LEARNED ABOUT EDUCATIONALLY BENEFICIAL ACHIEVEMENT TESTS

W. James Popham, IOX Assessment Associates and University of California, Los Angeles, USA

Introduction

During the year in which ECER 1992 takes place, it will have been three full decades since Robert Glaser and David Klaus first used the phrase ``criterion–referenced measurement" in a collection of essays edited by Robert Gagne. Thereafter, it took almost a half a decade for educational measurement specialists to become interested in the conceptualization of assessment that was embodied in the phrase ``criterion–referenced." For the past 25 years, however, the impact of a criterion–referenced assessment approach has been substantial. During those 25 years, I have been active in a variety of U.S. test–development efforts. The organization I head, IOX Assessment Associates, has created statewide criterion–referenced measurement systems for a dozen U.S. states as well as numerous local school systems. Given such a substantial body of test–development work, it should be apparent that my IOX colleagues and I have committed our fair share of measurement mistakes. These mistakes, however, have led to the generation of more defensible ways to create achievement tests so those tests make a positive contribution to the quality of the educational process. The lessons we have learned during the past 25 years are presented here in the form of five guidelines.

Guideline 1: Educationally beneficial tests should measure genuinely defensible outcomes

In order for educational tests to have a positive impact on instruction, those tests need to deal with worthwhile knowledge or skills. This first guideline sounds suspiciously like a platitude that must be ritualistically endorsed lest the writer seem indifferent to the significance of test content. Unfortunately, as bitter experience has shown us, there are often strong pressures in our educational enterprise to trivialize what we test on high–stakes examinations. Fearing adverse evaluations, some educators have endorsed decisively lower level assessment targets.

I do not wish to suggest that all of the U.S. educators who have played a role in determining the content of educational accountability tests placed their own interests ahead of the interests of students. On the contrary, during the past 20 years I have listened to numerous teachers and administrators argue eloquently

for the inclusion of genuinely challenging content on high–stakes tests. However, in too many cases when a content–advisory committee of educators had to choose between a more–demanding or a less–demanding way of operationalizing a skill, I've watched the vote go in favor of the less–demanding option. No committee members ever voiced the view that the less–demanding content would be easier to teach, but I'll wager that many committee members knew exactly what was going on.

How, then, do educators go about accomplishing Guideline 1 in a practical manner? My answer is quite straightforward. I think that it is imperative to establish a nonpartisan content–review group to consider the potential content to be included in a high–stakes tests. This does not mean that practicing educators, namely, teachers and administrators, should not play an important role in content–determination decisions. However, after such educatorshave finished their work, an independent group of non–educators should review the proposed content to see whether the curricular aspirations embedded in the proposed content are sufficiently meritorious.

Guideline 2: An educationally beneficial test should be focused on an instructionally manageable number of assessment targets

An instructionally manageable number of assessment targets consists of that number of assessment foci that a typical teacher will be influenced by in making instructional decisions. Usually, that number is around a half dozen or so. With many more assessment targets than that, a teacher tends to disregard evidence of students' achievements simply because there are too many things to which the teacher must attend.

The trick involved in implementing this guideline is to coalesce as many small–scope assessment targets as is instructionally reasonable so that teachers can be guided by a modest number of broad assessment targets that are still measurable. It is imperative as we collapse smaller skills or bodies of knowledge into larger, yet still instructionally sensible targets, that we retain measurability. If we move so far in our attempt to create broad–scope assessment targets that we have obscured the instructional essence of what is being measured, we've moved too far.

Guideline 3: Norm–referenced tests should not be used to support the instructional process

Because norm–referenced achievement tests are developed and distributed in the U.S. by large, for–profit corporations, it is not surprising that corporate shareholders would like their companies to make money by setting oodles of tests. But, curricular preferences in different geographic locales of the U.S. are often quite dissimilar. What is instructionally emphasized in the Boise, Idaho language arts program may be substantially different from what is emphasized in the Miami, Florida language arts program. Because of this widespread curricular diversity, the publishers of norm–referenced achievement tests dare not spell out with detailed precision just what their tests are measuring. Clearly recognized mismatches between what was being taught and what was being tested would lead to markedly reduced sales of norm–referenced achievement tests. Thus, chiefly for profit motives, publishers of norm–referenced achievement tests describe their wares very generally. General descriptions of what's being tested

allow prospective purchasers to assume that what the achievement test assesses will be essentially similar to what is being taught in local schools. Typically, this is an unwarranted assumption. Unfortunately, many of the mismatches between what's tested and what's taught are completely unrecognized by local educators who often derive at least a partial idea of how well they're teaching on the basis of their students' norm–referenced test performances.

A second shortcoming of norm–referenced achievement tests, from an instructional perspective, is that they provide insufficiently clear descriptions of what they measure. These inadequate descriptions do not permit teachers to design truly on–target instructional sequences because the teachers really don't have a clear idea of what's being assessed.

A final instructional shortcoming of norm–referenced achievement tests has been brought to our attention by John J. Cannell, a public spirited U.S. physician who, having become interested in the accuracy of test score interpretations, published a widely circulated monograph in which he revealed that almost all U.S. users of norm–referenced achievementtests were reporting that their students' performances were above average. Pointing out the inherent contradiction in the notion that everybody scores above the average, Cannell delivered a blistering attack on test publishers as well as public school educators regarding the credibility of norm–referenced test–score interpretations. As a consequence, many American citizens now believe that norm–referenced test results are not accurate because too many teachers have "taught to the test."

There is a role for norm–referenced achievement tests in education. That role should be restricted, however, to fixed–quota settings in which the most able or least able examinees must be identified. For purposes of helping classroom teachers do a better instructional job, norm–referenced achievement tests are of limited utility. Are norm–referenced tests better than no tests at all for purposes of instruction? Yes, they are. But just barely.

Guideline 4: Criterion–referenced tests must be constructed so that they foster generalizable knowledge and skills

Rather than focusing teachers' attention exclusively on one–definition testing targets, with the near certainty that nontransferable mastery by students will follow, we must build our test–item specifications for high–stakes criterion–referenced tests so that those tests foster generalizable mastery rather than stifle it. The strategy I recommend to accomplish this aim is to create a form of test–item specifications in which we set forth what is eligible to be tested, that is, the different kinds of testing tactics that might be employed to assess students' mastery. The specifications should be peppered with numerous sample items illustrating a range of items that might appear on any form of the test. Although, on any given form of a test, we might sample only a few of the eligible assessment tactics, all eligible assessment tactics would be fair game for any test. As a consequence, teachers would have to ready their students for different ways of assessing a skill and, thereby, would be promoting a student's generalized mastery of content.

Guideline 5: Teachers should be given assistance in learning how to appropriately prepare students for high–stakes tests.

This fifth guideline centers on an issue that has surfaced only in the last several years, namely, the appropriateness of the way teachers prepare their students for high–stakes tests. A pivotal assumption on which my views about educational testing are based is that high–stakes tests will inevitably have a profound impact on the instructional process. Even though high–stakes educational tests were forced on U.S. educators by noneducators, it is in the best interest of educators to have students do well on such tests. Accordingly, whether we like it or not, high–stakes educational tests will modify how teachers teach. In recognition of that reality, Guideline 5 points out that because teachers' test–preparation practices are sometimes inappropriate, teachers should be given sufficient staff–development assistance so that they understand how to link teaching and testing properly.

ANSWER QUESTIONS OR TELL THE STORY?

Lena Holmberg, Högskolan i Kalmar, Kalmar, Sweden

Introduction

This is a paper about the construction of test material for assessment of reading proficiency in grade 5 within the framework of the Swedish National Evaluation. The opinion has been put forward, from time to time, that pupils should have opportunities to give free expression to the associations which texts arouse in them even when it comes to evaluation. In short, one might present the discussion so, that a more traditional method of developing and testing reading ability is set against a holistic method. The debate which both points of views have given rise to in the field of education has also affected evaluation.

One could claim that the reading of literature in the classroom, where time is not limited strictly, where teachers and pupils can widen their horizons by answering and asking questions, explaining the meaning of words, talking about what they have read etc. is something else than the reading of literature in a controlled way in a national sample of classes and where the results are collected in order to give a basis for a current description of pupils' reading proficiency.

On the other hand the criticism that pupils' own thoughts are not given expression at the time they are answering given questions to a text can be justified. However, the questions in our test battery were based on pupils' ideas which came from freely-worded responses during the construction work. Even so, there could be grounds for studying the problems before a new round of the National Evaluation is put into effect in 1992.

Method

Ten classes took part in the study, six in grade 2 and four in grade 5. All the pupils in the classes read a story, "The Class", which was part of the test battery for grades 2 and 5 in the National Evaluation in 1989, and they replied to the questions on it. The result of this test was used as a measure of the pupils' reading proficiency, and our sample could be compared with a national sample. Nine pupils were selected from the classes in grade 5; three were poor readers, three average and three good. The class teachers made the choice. In grade 2, three average and three good readers were chosen.

Half, that is 36, of the selected pupils read another text, "Graffiti", answered questions to it and afterwards told the story in their own words (Group 1, A+T). The other half read the story, then told it, and after this they answered the questions (Group 2, T+A).

The pupils' answers to the question on "The Class" were scored in accordance with the guidelines, which were used by hundreds of teachers in the first evaluation (1989). The scoring of the questions on the text "Graffiti" was made by the test administrator.
While the pupils were telling the story "Graffiti", they were recorded on tape. For pedagogical reasons it was considered essential that the pupils were allowed to finish telling the story, with or without help. This meant that the narrative in some cases turned into something lika a dialogue. The transcripts from the tapes, however, were marked so that only those events described by the pupils before intervention were considered in the analyses.
The study was regarded as an experiment with repeated measurement, where the results on "The Class" represented the measurement before and the results on "Graffiti" the measurement after.

Results
The two groups' initial reading proficiencies differed. Grupp 1 (A+T) started with lower scores on the text "The Class" and achieved better results, relatively speaking, when answering the questions on the text "Graffiti".
When the result from "Graffiti" was measured by the number of elements in the pupils' narratives the groups were even closer. Despite their poorer reading proficiency Group 1, who answered the questions first, reached almost the same result as Group 2, who told the story first.

Conclusions
The result indicates that especially poorer readers benefit from reading tests, where they answer well-constructed question, which can in fact direct the pupils to the understanding of a text.
The consequense for the continued assessment of reading in the National Evaluation is that reading proficiency can very well be measured by well-chosen texts from children's literature and well-designed questions referring to them.

LISKAL: A PUPIL MONITORING SYSTEM BASED UPON A NEW LONGITUDINAL MEASUREMENT METHODOLOGY

Johan H.L. Oud, Jan F.J. van Leeuwe, Cor A.J. Aarnoutse, & Marinus J.M. Voeten, Catholic University of Nijmegen, Nijmegen, The Netherlands

A topic of major interest in psychology and education is the development and implementation of reliable and valid procedures for estimating individual subjects' values on latent variables. When a longitudinal explanatory state–space model (as estimated by means of the LISREL program, e.g.) is available for the latent variables to be measured, the powerful Kalman filter can be used for that purpose. It is known to be optimal and usually leads to considerably improved latent value estimates in comparison to those from conventional cross–sectional measurement procedures. The state space model can also be used to predict future achievements and, in combination with the Kalman filter, to evaluate remedial measures taken.

The implementation of the Kalman filter in LISKAL–M1 and LISKAL–M2 made it possible to reduce with 30 to 50% the estimation error variance resulting from Bartlett's well–known cross–sectional factor score estimator (Oud & Mommers, 1988). In both versions of the pupil monitoring system LISKAL the latent achievement variables (Decoding Speed, Reading Comprehension and Spelling) are present as standardized variables in the form of z–scores, having mean 0 and standard deviation 1 at each point in time. Because by means of z–scores the pupil is compared with a normative group of pupils, the assessment is relative. However, teachers are often interested in absolute developmental curves. A pupil's relative decrease of achievement in comparison to a normative group, for example, does not exclude an absolute increase in ability for that same pupil. For evaluation purposes assessment of absolute change can be as important as of relative change. Over the group as a whole absolute developmental curves have means, which are possibly different from 0 and standard deviations which, for instance, increase or decrease over time.

The problem of estimating absolute developmental curves can also be solved by means of the Kalman filter (Oud, van den Bercken, & Essers, 1990). The LISREL model on which it has to be based is more complicated, however. It includes a structured means part with special identification problems. In particular, identification should guarantee that the latent variables keep the same content across time.

The estimation of absolute developmental curves will be implemented in the new LISKAL–M3 program. LISKAL–M3 will be based on the longitudinal research

project LONGI which is currently carried out at the University of Nijmegen with a nationwide sample of 3430 pupils. Other improvements made in LISKAL–M3 on the basis of LONGI are the following.

- LISKAL–M1 and LISKAL–M2 were restricted to reading skills: Decoding speed, Reading Comprehension, and Spelling. LISKAL–M3 will also cover numerical skills.
- LISKAL–M1 and LISKAL–M2 followed the pupils over three years over primary education (groups 3, 4, and 5 of Dutch basic education). LISKAL–M3 will encompass 6 years of primary education (groups 3 through 8 of Dutch basic education). A special technique is used in the LONGI project to estimate the 6 years longitudinal LISREL model on the basis of data collected over 2 years. This prevents the pupil monitoring system from becoming available only after 6 years and thus being obsolete at the time it is published.
- LISKAL–M1 and LISKAL–M2 required the teacher to take tests at specific dates. However, because of illness of pupil or teacher or because of other reasons this often turns out to be impossible or difficult to realize in practice. In LISKAL–M3 a special technique, transforming the discrete state space model in a continuous one, is used allowing the teacher to take each test at an arbitrary date in a specific period.

The improvements made in LISKAL–M3 and its resulting flexibility make it usable as a a general evaluative assessment device in regular as well as in special education.

References

Oud, J.H.L., & Mommers, M.J.C. (1988). Longitudinale computerondersteunde onderkenning van lees– en spellingsmoeilijkheden: Een toepasing van het Kalman filter in de onderwijspraktijk [Longitudinal computer assisted assessment of reading and spelling difficulties: An application of the Kalman filter in educational practice]. Tijdschrift voor Onderwijsresearch, 27, 31–49.

Oud, J.H., van den Bercken, J.H., & Essers, R.J. (1990). Longitudinal factor score estimation using the Kalman filter. Applied Psychological Measurement, 14, 395–418.

AN ADAPTIVE TEST OF DUTCH AS A SECOND LANGUAGE

U. Schuurs, & J. de Jong, National Institute for Educational Measurement (Cito), Arnhem, The Netherlands

Education of Dutch as a Second Language (DSL) is provided by a large variety of courses at different levels of proficiency. In order to make a good choice an efficient placement test is crucial. As the language proficiency of second language learners varies considerably, a placement test has to cover a very broad range of language proficiency. This makes it necessary to use a test with many questions at different levels, as implemented in the well known standardized placement test 'Instaptoets Anderstaligen'. The size of this test however is incompatible with the required efficiency; as a result of this dilemma at many schools only a minor part of this DSL–test is used in second language placement decisions.
In order to challenge this problem an adaptive language test has been constructed. Main characteristic of adaptive testing is that the difficulty of the questions depends on in–between scores, so that candidates are not to be confronted with many questions that are far above or far below their actual proficiency level. Several adaptive testing systems can be distinguished, the most elementary of which relies on the branching principle. In more sophisticated systems the students ability is estimated after each response, so that from the item bank an item can be selected that best fits with the ability level demonstrated thus far. Such systems however are not well suited for the hardware available in Dutch school systems. For that reason we have chosen a setup according to the more simple branching system.

General test structure

The general structure of our system can be described as a matrix with two dimensions, each of which has five levels. The first dimension contains five domains of language use (morphology; language perception at vocabulary level and at sentence level; language production at vocabulary level and at sentence level). At the second dimension five levels of item difficulty are distinguised. As a result we can differentiate per language domain and per level of difficulty.
In our system testees are confronted default with three items of average difficulty (starting level 3). On the count of items answered correctly, the next string of three items from the same language domain will be from a higher level (all three answers correct), the same level (two answers correct), one level down (only one answer correct), or even two levels down (no correct answers). As the next item string is selected from another domain, level changes are not obvious for the testee. The procedure described above is repeated three times for each domain,

so that the test session is ended after 45 questions. At the end of a test session the student gets an outprint of the items that were answered wrongly. In practice it takes about 25 minutes to take the test.

Objective setting
Main purpose of the test construction is to facilitate intake procedures. In secondary education test results may also help decide which students are qualified for remedial education with a focus on the second languageacquisition problems. In business context the test results may help answering the question whether an employee needs training in the Dutch language.
But the test results also provide diagnostic information. At the end of a test session results are reported for each of five dimensions seperately. This enables teachers to decide on the basis of what skills they will emphasize in courses at different levels.
The test is not –as had been strongly recommended by Laurier & Des Brisay (1990)– a program-based language test, attuned to specific curricula. In the Netherlands most of the second language learning programmes are rather ecclectical. At elementary levels the test focusses on grammatical aspects of language proficiency, at the higher levels more communicative items are presented. Serious problems in construction were caused by the interface: as much research indicates that educational programmes are confusing as a result of too much information on the screen, we decided to incorporate items only if they did not exceed 3 lines (about 20 words, question and alternatives not included).

Validation procedure
Items were pretested with a group of second language learners, but also with Dutch students: the assumption was, that especially information was needed about differences in linguistic competence of Dutch and ethnic students (cf. de Jong 1991). As a second advantage the levels of difficulty can be described in terms of the Dutch school system. This seems a legal procedure as most problems in the second language acquisition are intralingual, that is independent of the first language (Verhoeven 1987).
In the validation procedure we used stratified sampling: all items of adjacent levels of difficulty were pretested at different levels of the Dutch school system. The assumption was that first language acquisition process is completed at the age of 15; for this reason we tested level 5 at students of grade 11. As regards to the lowest level we hypothized the language proficiency of 6th graders could function as a criterion. The design of the pretest is visualized in Table 1.

All students were confronted with items of all dimensions, but restricted to two adjacent levels of difficulty (and 6th graders to one level only). As an anchor each student got 10 more items, derived from the standardized 'Instaptoets Anderstaligen'.

Table 1
Design for the pretest

level of difficulty	n	1	2	3	4	5
grade 6	200					
grade 8	150					
grade 9 (MAVO)	150					
grade 10 (HAVO)	150					
grade 11 (VWO)	200					
Total n per level		350	300	300	350	200

Needed additional research

As the pretest was presented in a paper–and–pencil version, we have as yet no information on students attitudes towards the use of the computer as a test device. Other research, however, suggests that use of a computer adaptive test is preferred by students, possibly because it may be less threatening and less items are presented that are too difficult or too easy.

A serious disadvantage might be, that an adaptive test consists of fewer items so that test results are less reliable. On the other hand, some experiments suggest that adaptive tests of extreme short lenght may still be suitable as a screening test. For instance McBride (1990) compared eligibility determinations using a traditional pencil–and–paper test and an adaptive test consisting of 5 items. Determinations were in agreement in over 80% of 1500 cases.

As test results are based upon rather few items, all items have to be highly reliable, and have to be scaled. From our original pool of 300 items we have to select the 225 most reliable items. As a second step we have to scale the items in a calibration procedure. As the software is concerned we have to avoid unjustified rejection of answers, caused by incorrect spelling as regards to open questions. For this reason incorrect spelling alternatives that are frequently used by speakers of Dutch as a Second Language will be imported in the system.

References

Cito (1988). Instaptoets Anderstaligen. Arnhem (Cito), The Netherlands.

De Jong, J.H.A.L. (1991). Defining a variable of foreign language ability / An application of Item Response Theory. Dissertation Un. Twente.

Laurier, M., Des Brisay, M. (1990). An Integrated Approach to Developing Small-Scale Standardized Tests. Eric Documents (ED326052).

McBride, J. (1990). A computerized Adaptive Mathematics Screening Test: A Pilot Study. Eric Documents (ED 316554).

Verhoeven, L. (1987). Ethnic minority children acquiring literacy. Dissertation Un. Brabant.

SIMULTANEOUS OPTIMIZATION OF CLASSIFICATION DECISIONS FOLLOWED BY A MASTERY DECISION

H. J. Vos, Department of Education, University of Twente, Enschede, The Netherlands

Introduction

Decision problems in educational and psychological testing can be classified in many ways. An elegant typology of test-based decisions has been given in van der Linden (1985, 1990). Each type of decision making in this typology can be viewed as a specific configuration of three basic elements, namely a test, a treatment, and a criterion. In general, the following four elementary types of decisions can be distinguished: selection, mastery, placement, and classification. The four elementary decisions can also be considered in combination with each other. One combination, for instance, is test-based decision making in Individualized Study Systems (ISSs) (Vos, 1990, 1992; Vos & van der Linden, 1987). The purpose of this paper is to explore the simultaneous optimization of combinations of elementary decisions. To illustrate the approach, a classification (in case of two treatments) and a mastery decision are combined into a decision network and it is shown how so-called weak and strong optimal cutting points can be derived, i.e., decisions which explicitly take and do not take prior achievement on the classification test into account. Two main advantages can be identified of the simultaneous approach comparing to optimizing these decisions separately. First, decisions to be made later in the decision network can already be taken into account. Secondly, more realistic utility structures can be handled effectively by the simultaneous approach.

Methods and techniques

An appropriate framework for dealing with decision problems such as above is Bayesian theory (e.g., Lindgren, 1978). There are two fundamental elements in a Bayesian procedure: probabilities and utilities. In this paper, an additive utility structure is adopted in which the form associated with each separate decision is assumed to be a threshold function. The optimal procedure as prescribed by Bayesian decision theory is to look for a decision rule that maximizes expected utility. To derive the optimal monotone rules, methods of numerical optimization have been used.

Data sources and results

To illustrate the optimal decision procedures presented in this paper, a case with empirical data is given coming from a well-known problem in the Netherlands: the assignment of students to appropriate continuation schools at the end of the elementary school (i.e., at grade 8). In general, the transition process from the elementary to the secondary school is based on the teacher's recommendation, parents' preference, and an achievement test administered in February or March at grade 8. The Dutch National Institute of Educational Measurement (CITO) prepares annually an achievement test, which is used by most elementary schools for this purpose. In addition, on the basis of the grade-point average, it is decided whether or not a pupil finishes the first year of secondary school succesfully. This means that the problem can be characterized as a combined classification-mastery decision problem. It turned out that with the simultaneous approach students were sooner assigned to higher types of secondary education than with the separate approach. In particular for a low weight for the utility associated with the separate classification decision, pupils were assigned much sooner to higher types of education. On the other hand, the optimal mastery rules did not yield large differences. Furthermore, the overall expected utility both for the simultaneous and separate approach was computed to investigate whether there was a gain in overall expected utility for the simultaneous approach compared to the separate approach. It turned out, however, that the separate overall expected utility was larger than the simultaneous overall expected utility. This unexpected result could be explained by the fact that the so-called monotonicity conditions were not satisfied. Finally, it could be shown that the overall expected utility for the weak monotone solution was slightly larger than for the strong monotone solution. This result was in accordance with a Lemma shown in the paper.

Conclusion and discussion

Although the simultaneous approach did not yield a larger overall expected utility than the separate approach, due to the fact that a local maximum instead of an absolute maximum was found, the simultaneous model still seems to be a promising new approach to derive optimal rules for combinations of elementary decisions. More research is needed in other optimization methods than Newton-Raphsons's iterative procedures, which was used in this paper. For instance, a first-order method such as steepest descent or quasi-Newton techniques may be applied. A final remark is appropriate. The models presented in this paper have a larger scope than selecting optimal continuation schools. For instance, the qclassification-mastery decision problem may be important in classification-mastery decision problem may be important in classification of students in ISSs with tracks at different levels followed by a mastery test at the end of the tracks. Another example may be found in the area of psychotherapy in which patients have to be assigned to the most promising therapy followed by an end-of-therapy test.

References

Lindgren, B.W. (1976). Statistical theory (3rd ed.). New York: Macmillan.

Linden, W.J. van der (1985). Decision theory in educational research and testing. In T. Husen & T.N. Postlethwaite (Eds.), International encyclopedia of education: Research and Studies, pp. 1328-1333. Oxford: Pergamon Press.

Linden, W.J., van der (1990). Applications of decision theory to test-based decisions. In: R.K. Hambleton & J.N. Zaal (Eds.), New developments in testing: Theory and applications, pp. 129-155. Boston: Kluwer.

Vos, H.J. (1990). Simultaneous optimization of decisions using a linear utility function. Journal of Educational Statistics, 15, 309-340.

Vos, H.J. (1992). Simultaneous optimization of the aptitude treatment interaction decision problem with mastery scores. In M. Wilson (Ed.), Objective Measurement: Theory into practice, to appear. Norwood, New Yersey: Ablex Publishing Corporation.

Vos, H.J. & van der Linden, W.J. (1987). Designing optimal rules for instructional decision making in CAI systems. In J. Moonen & T. Plomp (Eds.), Developments in educational software and courseware, pp. 291-298. Oxford: Pergamon Press.

A DECISION SUPPORT SYSTEM FOR THE EVALUATION OF COURSEWARE QUALITY

Cees Volman, & Italo De Diana, University of Twente, Enschede, The Netherlands

A description of the problem

Daily courseware use may demand "tailor made" courseware product adaptations at "programming level" to be implemented the teacher. Such adaptations are usually carried out based upon the perceived pragmatic quality of courseware: the quality of courseware as becomes apparent as a result of its use. The Process of Product Quality Analysis and Improvement (for short PQAI) (De Diana, 1988) requires the teacher to make decisions in respect to:

- a (measurable) quality concept;
- a quality analysis strategy;
- a product improvement selection strategy;
- an adaptation implementation strategy (Volman, 1990).

The PQAI Process is considered a too complex, intransparant and difficult process for teachers (but not only for them,Hall & Hord, 1987). To complicate matters even more, teachers differ in respect to affinity and experience with courseware use and optimisation. Yet the demand for product adaptation clearly is there.

A Decision Support System (DSS) could support the PQAI process and thus could promote the pragmatic quality of courseware products. This paper describes such a DSS system that has been developed at the University of Twente.

The Decision Support System

Decision Support Systems are comprised of techniques to compensate for human cognitive deficiencies regarding decision making problems (Zachary, 1988). Related to the PQAI process, four classes of problems have been identified (Volman, 1990):

- data storage, maintenance and retrieval problems;
- information interpretation problems;
- reasoning problems, and
- problem representation problems.

The DSS that has been developed to support the PQAI process helps teachers in each of these four categories of problem solving. Problem class one is supported by the implementation fo a data file management system. Problem class two is supported by the implementation of a data analysis function, making it possible to compare recorded values with presepecified criterion values. Problem class three

is supported by the implementation of a function that "translates" analysis results into "appropriate" adaptation advise. Problem class four is supported by visualising techniques for:

- showing the hierarchical and modular structure of courseware;
- showing the analysis results that do not match prespecified criterion values and
- showing selected "appropriate" adaptation advise.

The developed DSS supports both the *qualitative* and *quantitative* side of the PQAI process. Regarding the *qualitative* analysis process, the DSS uses the Component Display Theory from M.D. Merrill (Reigeluth, 1983). The qualitative analysis processing is implemented as the matching between the courseware CDT code sequences and the CDT code sequences as suggested by the CDT.
Regarding the *quantitative* analysis process, the DSS uses a "global versus detailed" analysis control mechanism. This control mechanism is used to account for differences in teacher perceptions regarding the PQAI process. The control mechanism is designed as a "fill out the form" procedure, to be carried out once for each level in a hierarchic courseware product. Each form will be used for all the elements at a typical courseware level.
On each form, a teacher has to specify several options or values. Firstly, a choice is to be made between primary group directed or primary learner directed processing. This means that a teacher wants the PQAI process to be continued for individual learners only if the analysis results for a group of learners is satisfactory. In this sense, the "group" option indicates the global PQAI approach, the "learner" option the detailed PQAI approach. Secondly, a teacher must specify "depth" or "breath" to indicate the direction of the PQAI process. If "breath" is selected, then the DSS processes all elements on one typical level before descending to the next level. If "depth" is selected, then the DSS descends to the next level after having completed the analysis for a typical element. If for all forms the "breath" option is used then a global PQAI approach is reflected. If "depth" is used then the detailed PQAI approach is supported. Thirdly, a teacher must specify one or more variables and the according criterion intervals. The variable(s) can be used for both group and individual learner processing. Finally, the teacher must specify one or more IF..THEN..ELSE..rules. The conditional part of a rule is used to match the (combination of) recorded value(s) for the variable(s) with the prespecified interval(s). If a teacher defines only a few simple rules, the global PQAI approach is reflected. If many complex rules are used, the detailed PQAI process is supported.
Both the qualitative and the quantitative PQAI ananlysis strategy of the DSS enable a teacher to conduct the PQAI in a systematic and methodic way without burdening him with the necessity to have detailed knowledge regarding an instructional design theory or authoring tools.

Usability of the DSS explored

The usability of the DSS has been explored by teachers in a three-phase procedure (Volman, 1990): 1) introduction to the DSS, 2) demonstration of the DSS by the designer, and 3) use of the DSS by the teachers. The tray-out is conducted within the context of the "PRONTO-USE" research project.

Conclusions

The main conclusions of the DSS usability exploration were:

- the DSS user interface matches the evaluation processing structure in daily practice;
- stating the PQAI process explicitly in terms of the quantitative quality definitions in order to allow for automated DS requires learning time but is feasible;
- a valid qualitative and quantitative "improvement" knowledge base and an on-line help system arae to be regarded as vital parts of the DSS.

This exploration has indicated:

- a DSS as described to be a valuable teacher aid in the PQAI process;
- the concept of "pragmatic" quality is shown to be "measurable" in a systematic and methodic manner;
- further research concerning the PQAI process is necessary in order to warrant the validity of the DSS advise.

References

De Diana, I. (1988). Het EDUC-systeem: aspecten van een methodologie, ontwikkelingsmethode en instrumentatie voor tutoriael COO. Dissertation. Enschede: Universiteit Twente.

Hall, G.E., & Hord, S.M. (1987). Change is the schools: facilitating the process. New York: State University of New York Press.

Reigeluth, C.M. (Ed.) (1983). Instructional- design theories and models: An overview of teh current status. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Volman, C.J.A.M. (1990). Kwaliteitsbepaling van interactieve onderwijs-leermiddelen: ontwikkeling van een prototypisch beslissingsondersteundend programmatuursysteem. Doctoraalscriptie. Enschede: Universiteit Twente.

ADAPTATION AND STANDARDIZATION OF THE TEACHER VERSION OF THE CHILD BEHAVIOR PROFILE: TURKISH GIRLS AGED 7 - 12

Füsun Akkök, & Petek Askar, Faculty of Education, Middle East Technical University, Ankara, Turkey

Introduction

This paper reports on the adaptation and the standardization of the Teachers' Report Form(TRF) of the Child Behavior Checklist CL) and profile for girls aged 6-11 to the Turkish girls aged 7-12. The Child Behavior Checklist, which has been developed by Achenbach (1978) is an assessment instrument designed to obtain the parents' reports of their children and to provide a standardized frame of a child's behavior related to behavioral and adaptive competencies.

The TRF was aimed to obtain the reports of the teachers in connection with their students' problems and adaptive functioning in a standard format. It has been standardized separately for boys and girls at ages, 6-11 and 12-16, to reflect the age and sex differences in the prevalence and of behavioral patterns. The norms of each scoring profile has been established by a representative sampling of nonreferred children in order to provide comparisons of data or individual children with the baselines for normal peers.

In this paper, we present a Turkish adaptation and standardization of the TRF and describe the standardized profile obtained. We have standardized the scores based on teachers' ratings of girls aged 7-12, since in Turkey, elementary school age is 7-12. In two independent studies, the percentage of referrals to clinics was found to range from 64.5% to 35.5%; and from 61.5% to 38.5%, for boys and girls respectively (Epir,1974 and Sonuvar et al.,1982). Due to these findings, we first carried on a study on the adaptation and standardization of the TRF for boys aged 7-12. Data for that study were obtained from 48 referred and 294 nonreferred boys. Comparison of referred and and nonreferred samples showed significant differences on all behavior problem scales, except the "Anxious" scale. The median of the Cronbach internal consistency of the scale scores was 0.75 (Akkök and Askar,1989). In order to have a scale adaptable for all elementary school population, we continued our research on Turkish girls aged 7-12.

Method

Subjects

The sample consisted of 231 nonreferred girls. While drawing randomly the normative sample of 231 girls, special attention was given to include equal

762

number of girls from each age group between 7 and 12. The normative sample consisted of children who had received no mental health services and was used as a criterion group for obtaining evidence of the validity of the TRF. TRFs were obtained from teachers of regular classes in public and private schools in Ankara. Elementary school teachers with different number of years of professional experience, rated the students whom they had known for at least two months. The familiarity of the teachers with the students depend on the grade of the students, since in Turkey, the same teacher usually teaches all through the elementary school years. In this study, three of the schools were from high SES, 8 from middle and 4 from low SES areas. The thirty-eight referrals came from the 3 mental health clinics in Ankara specialized for children. Data were collected over a period of 3 months.

Procedure

In the Turkish form of the TRF as in the original form, the first page covers demographic information including the child's age, sex, grade and parents' occupations. It also includes questions about the child's academic performance and clinically useful background information.

The four questions on the second page regarding the adaptive behavior functioning have been excluded from the Turkish form, since there are no available standard achievement, IQ, readiness and aptitude tests which are used during elementary school years. Therefore, these questions seemed irrelevant and were omitted from the Teacher's report form. As in the original form, space was provided for teachers to report on the child's work, behavior or potential.

The third and the fourth pages of the TRF included 118 Behavior Problem items. In the Turkish form,item 105 " Uses alcohol and drugs" was omitted, since it seemed irrelevant to the Turkish children of that age group (7-12).

Results

Factor Matrix

Principal-component analysis was performed on the eight behavior problem scales for the 38 clinically referred girls aged 7-12 to see whether the scales could be grouped as internalizing and externalizing. The first factor includes Unpopular, Self Destructive Nervous-Overactive and Aggressive Scales. The second factor includes Anxious, Social Withdrawal, Depressed and Unpopular. The third factor includes Social Withdrawal and

Inattentive. Depressed and Unpopular both have high loadings on factor 1 and factor 2.

Father's Occupation and Referred Versus Nonreferred Comparisons Behavior Problem Scales

Unweighed means analysis of variance (ANOVA) was used to evaluate differences in behavior problem scales due to Father's Occupation and clinical status for the 7-12 girls. Separate 3x2 ANOVA's reveal significant effects of clinical status for all eight behavior problem scales. The F values ranged from 10.51 to 88.40, (all $p<0.01$). For each scale, clinically referred girls obtained higher scores than nonreferred girls. Father's occupation was not involved in any significant main effects for either group of girls. However, significant Clinical Status x Father's

Occupation interactions were detected for the Anxious Scale $F(1,205)=4.84$, $p<0.05$, and Unpopular Scale $F(1,205)=4.67$, $p<0.05$.

Internalizing, Externalizing and Total Score

Separate 2x2 ANOVA's revealed significant effects, ($p<0.001$) of clinical status for Internalizing ($F=31.927$), Externalizing ($F=121.960$) and total behavior problem score ($F=275.4$). For each scale clinically referred girls obtained higher scores than nonreferred girls. No interaction effects were significant in these analysis. However the main effect of Father's Occupation $F(1,205)= 6.186$, $p<0.002$, on externalizing scores was significant.

The median of reliability estimates of scale scores is 0.84, and total score is 0.92. The percentiles and normalized T scores of the profile were derived from data provided by the teachers of the 231 girls, aged 7-12, who were randomly selected from regular classrooms.

Conclusion and Discussion

Based on the results, it can be concluded that TRF can be satisfactorily used for the Turkish girls 7-12. With the previous adaptation and standardization of the boys 7-12, the Turkish version of the TRF can be a good source for clinicians to learn about referred elementary school girls' and boys' behavior problems.

References

Achenbach,T.M.(1978) The child behavior profile:Boys aged 6–11. Journal of Consulting and Clinical Psychology,6,478–488.

Akkök,F. & Askar,P.(1989) The adaptation and standardization of the teacher version of the child behavior profile: Turkish boys aged 7–12.International Journal of Psychology,54,129-136.

Epir, S.(1974) Survey of referred children to a clinic. Paper presented at the 10th annual convention of the Psychiatric and Neurological Sciences Association, Ankara, Turkey, August.

Sonuvar, B.,F. Oktem, A.Yorukoglu and S.Akyildiz.(1982) The demographic characteristics of the referred children to Hacettepe Child Clinic in two years. Bulletin of the Turkish Psychological Association, 13, 33–39.

NEUROPHYSIOLOGICAL BASIS OF COGNITIVE ABILITIES

A.N.Lebedev, A.V.Markina, & I.V. Maltseva, Institute of Psychology AS USSR, Moscow, Russia

Abstract

Using our own as well as known neurophysiological data of M.N. Livanov we supposed that the diversity of phases and periodes of rhythmic brain processes limits the span of long– and short–term memory. We derived some mathematical equations of memory span from these neurophysiological premises. Two parameters of electroencephalogram were included into the formulas, namely, the peak frequency in alpha range of power spectra and the width of spectral "mountain" in this range. The latter is interrelated with the duration of single alpha spndles. To test our model of cognitive operations experimentally, we have used original programs of processing EEG and psychological data, specifically we have used the interactive procedure for memory span measurment. About one hundred subjects were tested. We conclud that the EEG parameters can be used for testing of cognitive abilities in human brain.

763

GUESSING-PRONENESS, BEHAVIOURAL AND COGNITIVE IMPULSIVITY: SEX DIFFERENCES

D. Bratko, K. Squarcy, P. Zarevski, University of Zagreb, Croatia; P. Kovacevic, & S. Logar-Djuric, University of Belgrade, Yugoslavia

In addition to knowledge, the students' results in objective tests are influenced by their response styles. These styles seem to be dependent on some personality traits, cognitive styles and test instructions. Guessing-proneness is one of the response styles in solving objective tests. It refers to risk-taking in test situations, which may be related to risk-taking activities in real-life situations.
Numerous studies reported great interindividual differences in guessing-proneness and complex relations between guessing-proneness and certain personality traits (Choppin, 1975; Hopkins,1972; Swineford and Miller, 1953; Zarevski and Rijavec, 1990). One of the personality traits that could be highly related to guessing-proneness is impulsivity , but it is necessary to make a difference between two types of impulsivity: behavioural and cognitive. Behavioural impulsivity concerns the level of the subjects' self-control, i.e. deficiencies in behavioural inhibition. Cognitive impulsivity is one of the best operationalized cognitive styles, concerning the speed of decision making in ambiguous conditions. It is interesting to point out that most of the studies reveal no relation between cognitive and behavioural impulsivity.

Purpose

The purpose of this study is to examine sex differences in the relationship between GP and behaviourally and cognitively defined impulsivity.

Variables

Guessing-proneness is operationalized as a number of answers to questions with no possible correct solution. This is the most direct measure of guessing-proneness. The psychometric validity of this procedure has been proven in several studies (Kovacevic and al.,1991; Zarevski and Rijavec, 1990). We used the General Information Test (GIT; Zarevski,1988) consisting of 66 "straight" items and 24 questions with no possible correct answer. GIT provided measures for the following variables: number of answers to questions with no correct answer (GUESS); number of correct answers (CORR); number of incorrect answers (FALSE); and number of skipped questions (SKIP).
Reflexivity-impulsivity was measured by use of the Gibson Spiral Maze Test (GSMT) (Gibson, 1964). Subjects' results on R-I dimension were calculated in two

ways: the subjects were split according to median of the variables "number of errors" (ER) and "time" (T) into four groups: reflexive, impulsive, fast and accurate, slow and inaccurate; as an individual ER/T ratio.
Two measures of behavioural impulsivity are used: Impulsivity Scale (Eysenck and Eysenck, 1977) consisting of 3 subscales: Narrow Impulsivity Subscale (NARR); Risk Taking Subscale (RISK); Nonplanning Subscale (NONPL); and Impulsivity Subscale (IMP) of the EPI (Eysenck and Eysenck, 1964). The complete EPI was administered, so the results on dimensions: Sociability (SOC), Neuroticism (N) and Dissimulation tendency (L) were also included.

Subjects
The sample consisted of 212 1st and 2nd year college students (122 girls and 90 boys), aged 15-16 yrs.

Results and discussion
ANOVA (Table 1) yielded statistically significant sex differences on following variables: number of correct answers on GIT; number of skipped questions; neuroticism; cognitive impulsivity (ER/T). Resulting sex differences are in accordance with previous findings which showed higher guessing-proneness for males (Choppin, 1975), and a tendency of skipping questions for females (Zarevski, 1988). Results indicate higher cognitive impulsivity (ER/T), and lower N for males. It is interesting to point out that no sex differences were found on behavioural impulsivity measures.
One-way ANOVA in the female sample, using cognitive impulsivity as an independent variable (defined by median) resulted in a significant F-ratio ($F = 5,34$; $p< 0.05$). Reflexive women answered on $M = 10,02$ questions with no correct solution ($SD = 6,12$), while impulsive women answered $M = 13,23$ ($SD = 6.34$) of them. In males such a difference was not found, although they are more impulsive and guess more than women.
In regression analysis the number of answers on questions with no correct solution was used as a criterion variable. Predictors were the measures of cognitive and behavioural impulsivity, along with the measures of sociability, neuroticism and dissimulation tendency. Results of the stepwise regression analysis for male and female samples are shown in Table 2 and Table 3, respectively. Regression analysis showed that the present set of predictors explains the guessing-proneness in the male sample better than in female. Stepwise regression analysis resulted in multiple correlation of $R = .547$, and of $R = .336$ for males and females, respectively.
The structure of predictor variables differs in male and female samples. In males, the measures of behavioural impulsivity proved to be the most predictive. Guessing prone males are low scorers on Narrow Impulsivity Subscale an Impulsivity Subscale of the EPI. They also seem to be risk-takers, prone to the dissimulation and nonplanning. On the other side, within the female sample the measure of cognitive impulsivity proved to be the most predictive. For the females, only 3 variables remained in the final equation. Guessing prone women seem to be more cognitively impulsive, more likely to plan and score higher on Narrow Impulsivity Subscale.

Conclusion

Obtained findings reveal a variety of sex differences in response style on objective tests, as related to some personality traits and cognitive style.

References

Choppin, B.(1975). Guessing the answer on objective tests. British Journal of Educational Psychology, 45, 206-213.

Gibson, H.B.(1964). The Spiral Maze: a psychomotor test with implications for the study of delinquency. British Journal of Psychology, 55, 219-225.

Hopkins, K.D.(1972). Response styles, chance, and other irrelevant sources of variation in test performance. In: Bracht, G.H., Hopkins, K.D., & Stanley, J.C.(Eds.), Perspectives in Educational and Psychological Measurement, Prentice-Hall, Engelwood Cliffs, New Jersey.

Kovacevic, P., Logar-Djuric, S., & Zarevski, P.(1991). Tendency to answer False Questions in Objective and Attitude Tests. 7th European Meeting of the Psychometric Society, Trier. 81-90.

Swineford, F., & Miller, P.M.(1953). Effects of directions regarding guessing on item statistic on a multiple-choice vocabulary test. Journal of Educational Psychology, 44, 129- 139.

Zarevski, P., & Rijavec, M.(1990). Personality correlates of guessing proneness on objective tests. Revija za psihologiju, 20, 27-33.

Appendix

Table 1

Means and standard deviations of male and female subjects and univariate ANOVA F-tests (* p<.05; ** p<.01)

	MALE		FEMALE		
Variables	M	SD	M	SD	F
TRUE	48.10	8.61	46.35	9.39	1.573
FALSE	10.44	8.03	9.48	8.25	.579
SKIP	7.47	6.73	9.86	7.06	5.030*
GUESS	13.83	6.86	11.61	6.44	4.781*
IMP	5.49	1.64	5.10	1.86	2.118
SOC	7.61	2.61	7.68	2.52	.038
E	14.34	3.69	14.02	3.98	.287
N	11.16	4.11	13.17	4.65	8.718**
L	2.93	2.05	2.63	1.65	1.178
RISK	2.71	2.01	7.27	2.09	.034
NARR	7.24	2.59	7.26	3.23	.002
NOPL	7.35	2.09	8.00	2.27	3.664
ER/Tx100	35.98	34.33	27.21	19.32	4.642*

Table 2

Stepwise regression of guessing proneness in males

Variables	Beta	Correl.	Part. corr.	T	p
NARR	-.479	-.264	-.419	-4.037	.000
RISK	.456	.209	.377	3.633	.001
NOPL	.307	.125	.260	2.500	.015
L	.224	.149	.213	2.048	.045
IMP	-.209	-.207	-.176	-1.697	.095

R = .547; R Square = .299; Adjusted R Squared = .245.

Table 3

Stepwise regression of guessing proneness in females

Variables	Beta	Correl.	Part. corr.	T	p
ER/T	.239	.216	.232	2.464	.015
NOPL	-.254	-.178	-.228	-2.415	.018
NARR	.237	.069	.208	2.211	.029

R = .336; R Squared = .113; Adjusted R Squared = .087.

EVALUATION OF TRAINING PROGRAMMES: THE STATE OF THE ART

J. de Koning, Netherlands Economic Institute, Department of Labour-Market and Education, The Netherlands

The paper deals with the 'state of the art' of evaluating training programmes for the unemployed. Such evaluating research centres around two questions:

- How far can a person improve his or her position on the labour market by participating in a training scheme?
- How far do training schemes stimulate economic development?

The emphasis is on the external return on training. However, its internal return, which depends on the duration and results of training, is also considered important, for a low internal return raises the cost of training schemes and reduces the external return.

In the past twenty years, evaluation methods and techniques have been developed and tested, and much empirical research has been carried out, mostly to measure the improvement in labour-market positions and much less to assess effects on economic development. The paper treats such questions as: what has been achieved, what gaps of knowledge remain, are the available methods and techniques satisfactory, and what further research is needed? Another question considered is whether methods and techniques evolved for the evaluation of training programmes are also applicable to other forms of education and training.

The paper strongly emphasises the need to assess net effects. Policy makers still tend to measure the success of training measures by the percentage of students who find a job after completing the course. However, there is no certainty that people find jobs because they have participated in a training course.

Several studies (among them some studies we did on Dutch training schemes) indicate that training is most effective for those groups among the unemployed who have a very low probability of finding a job. However this is often obscured by the fact that, although training is effective, the gross placement results for these groups remain relatively low.

Several methods have been developed to ascertain the net effect of schemes, the control-group method being one. The paper discusses several variants of that method. Both experimental and non-experimental designs are reviewed. The paper shows that even in the case of experimental design, the trainees and the controls are never fully comparable, because they have by definition different labour market "histories". Despite that fact, experimental designs must be strongly

770

preferred to non-experimental ones. Empirical studies, including research we did for the Netherlands, will serve as an illustration.
Particularly in American literature, the control-group approach is advertised as the most appropriate evaluation of training measures. That is not entirely justified as this method can only reveal how a certain scheme has affected the position of the participants on the labour market. To asses the effects on economic growth, due for instance to the elimination of bottlenecks on the labour market through training schemes, other methods are still waiting to be designed. Both theoretical and empirical research have challenged the claim made by human-capital theory that training more or less automatically increases productivity. The paper discusses the various positions and gives some leads on the development of appropriate methods for the assessment of the economic effects of training.

THE ASSESSMENT OF WRITTEN COMPOSITION

George Engelhard, Jr., Emory University, Atlanta, USA

The direct assessment of writing competence based on student compositions presents a number of measurement problems that cannot be systematically addressed within the framework of current test theories. Many of these measurement problems become compounded when the writing assessments are conducted within the context of large-scale state, national and international assessments. Ideally, the estimate of an individual's writing competence should be invariant over the particular raters, writing tasks, and domains that happen to be used in the assessment process. In practice, previous research has indicated that a variety of factors in addition to competence can contribute to the variability in observed composition scores. Some of the major factors are differences in rater severity, writing-task difficulty, domain difficulty when analytic scoring is used, examinee characteristics in addition to writing competence, and the structure of the rating scale.

One approach to the measurement problems encountered in direct writing assessments is to conduct a generalizability (G) study in order to identify the sources of error variance that contribute to the unreliability of the composition scores . Once the influences of various facets are identified, then steps can be taken to increase the reliability of the observed composition scores, such as increasing the number of raters. After a G study has been conducted and appropriate actions taken, there may still be significant sources of error variance in the operational assessment program. At this point, it may be possible to calibrate key facets, such as raters and writing tasks, within the framework of a multivariate latent variable model. If the facets can be reliably and validly calibrated, then adjustments can be made in order to achieve invariant estimates of writing competence.

The purpose of this study is to describe and illustrate the use of a Many-Faceted Rasch (FACETS) model that can provide a framework for obtaining invariant estimates of writing competence. The FACETS model is a multivariate extension of the Rasch model that can be used to calibrate multiple facets, such as raters, writing tasks and domains, that appear within the context of large-scale writing assessment programs.

Method

The data used in this study to illustrate the FACETS model is a random sample of 266 compositions written by 8th grade students as a part of a statewide assessment conducted in Georgia in 1990. These compositions were rated by 15

772

raters on 5 domains (content/organization, style, sentence formation, usage and mechanics) using a four category rating scale (1=inadequate, 2=minimal, 3=good, 4=very good). Each of the compositions was rated by two raters.
The measurement model underlying the current study is presented in Figure 1. The FACETS computer program (Linacre, 1988) was used to analyze the data, and Linacre (1989) should be consulted for further details of the FACETS model.

Results

Significant differences in rater severity were found, c (15) = 177.35, $p < .01$, with a high reliability of separation index ($r = .88$). Rater severity ranged from -1.36 to .98 logits. There was no evidence of significant rater misfit.
Observed raw scores ranged from 10 to 40 (M = 17.4, SD=6.1). On the logit scale, the Rasch estimates of writing ability ranged from -6.56 to 6.35. The differences between students were statistically significant, c (15) = 3554.41, $p < .01$,with a high reliability of separation index ($r = .93$).

Discussion

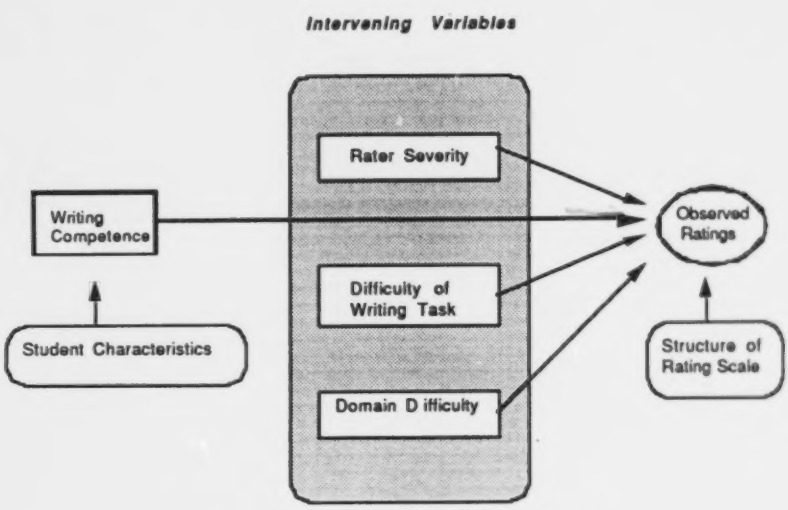
The results of this study suggest that there are significant differences in rater severity that influence the estimates of writing competence for these students. If adjustments are not made for differences in rater severity, then the estimates of writing competence will not be invariant over raters. For example, if a student is lucky and obtains an lenient rater, then he or she will have a higher estimate of writing competence; conversely, if the same student is unlucky and is rated by a severe rater, then a lower estimate of writing competence will be obtained for the same written composition. Large-scale writing assessments are too important to be dependent on the luck of the rater draw. The FACETS model offers a systematic framework for addressing measurement problems encountered in the large-scale assessment of writing. When the model fits the data, then objective and fair estimates of writing competence for individual students can be obtained that are invariant over particular raters and other aspects of the assessment process.

References

Linacre, J.M. (1989). Many-faceted Rasch Measurement. Chicago: MESA Press.
Linacre, J.M. (1989). FACETS: Computer Program for Many-Faceted Rasch Measurement. Chicago: MESA Press.

Figure 1

Measurement model for the assessment of writing competence



THE CONSTRUCTION OF A WIDE RANGE READING COMPREHNSION TEST FOR ADULTS

W. de Geus, O. W. H. Henneken, & P. Reitsma, Paedologisch Instituut Amsterdam, Duivendrecht, The Netherlands

Introduction

Recent disputes about adult functional literacy strongly suggest that one needs to consider carefully whether and how functional literacy differs from reading and writing abilities as taught in primary and secondary education. In order to be able to compare reading comprehension as measured by a traditional reading comprehension test with functional reading tasks, a traditional reading comprehension test had to be constructed for a young adult population. The purpose of the test is to indicate reading proficiency in a heterogeneous population that ranges from young adults still at secondary school to adults who already left formal education. The test should be able to reliably distinguish different ability levels and, in addition, it should be short and easy to administer. The latter was attained by using items in multiple choice format.

In order to avoid a theoretically complicated issue, namely what reading comprehension is, the reading comprehension test (RCT) is constructed to be closely analogous to the Dutch central exams for secondary education. There are roughly four levels in Dutch secondary education: lower vocational education (LBO), lower secondary education (MAVO), higher secondary education (HAVO) and preparatory scientific education (VWO). At the end of secondary education all students at a particular level of secondary education have to take a central examination, part of which is a test for reading comprehension. At LBO, MAVO and HAVO the central examination in the Dutch language consists of a reading comprehension test made up of several texts with questions. Our test must incorporate at least these 3 levels. In order to obtain content validity the RCT has to be similar in content and format to these tests. Construct validity is assumed to be accounted for by this content analogy (homology by analogy).

Construction

The construction has two phases in which texts were administered to secondary education pupils. Following the first phase tests were selected which can discriminate at each educational level separately. In the second phase these selected tests were combined in a single test that had to be applicable to all educational levels.

Construction of LBO and MAVO tests

From recent exams designed for the relevant levels we selected five texts that were both short and provided with suitable multiple choice questions. For three of these texts (one of which will become the anchor-test; see below) psychometric qualities as given by classical and IRT indices were determined using large national samples of pupils who had made these texts as an exam. Where necessary, adjustments were made to the item content and to the distractors. Badly discriminating or flawed items were deleted or rephrased and a few new distractors added. For the remaining two texts we judged the qualities of items guided by item writing prescriptions (De Groot & Van Naerssen, 1969; Mellenbergh, 1971; Van Berkel, 1984; Kok, Krieken & Luijten, 1986; Verhoeven & Beuk, 1986). Where judged necessary, adjustments were made to the item content and to the distractors.

Construction of HAVO test

To operationalize the HAVO level a different approach was needed. We could not use existing exams for this level, because the HAVO-exams measure reading comprehension by means of much longer texts with open-ended questions only. In order to meet the demands initially determined for the RCT the usual means of measurement for this level had to be adjusted. The characteristics of texts used in recent exams of higher secondary education were described by the mean text-length, mean sentence-length and mean number of difficult words per text. Two short texts were designed that were applicable to this level on the basis of these text characteristics (apart from text length). Open-ended questions were designed in such a way that the distribution of items, according to the categories of a classification scheme for reading comprehension items (Kreeft, 1985), was equal to the average distribution of items for this level. Two teachers judged both texts and questions to be suitable for higher secondary education and decided upon the correct answers. The two tests were administered to 134 pupils of the fourth and fifth grade of HAVO. Seriously flawed items were deleted and the resulting items were rewritten to fit the multiple choice format. From the wrong answers distractors for the multiple choice items were made. This resulted in two HAVO texts with 15 - 16 items.

Apart from the construction of six texts, two texts for each of the three educational levels, an anchor test was needed in phase 1 to check upon the operationalization of the three levels, to calibrate all items on a common scale and to estimate the proficiencies of all persons on the same scale. A text plus 13 items taken from a recent exam at the lower vocational level was chosen, based on IRT-analysis of a national sample of 2500 pupils.

The item-parameters were calibrated with the BILOG estimation procedure of Bayes' Modal Marginal A Posteriori assuming that 1) the population distribution of abilities is standard normal 2) the prior distribution (of parameter estimates) of the slopes of the ICC's is log-normal with mean 0 and s.d. 1.649 3) the prior distribution (of parameter estimates) of the item-difficulties is standard normal 4) the so-called 'pseudochance' parameter follows a bèta distribution with mean 0.25. Analyses (BILOG quasi-χ^2_{64}=503, p=.000, Mokken and LISREL unidimensionality checks, itemparameter-invariance checks) revealed that the 3-parameter-logistic-model fitted the data reasonably well, indicating that (part of)

this text could serve as an anchor test. The 5 best discriminating items were left unaltered to make linking of parameters possible. Where necessary, adjustments were made to the item content and to the distractors of badly discriminating items.

Phase 1

Thus far we constructed a total of 7 texts: one anchor text and two texts suitable for each of the three educational levels. These texts were administered to 650 fourth grade students, distributed over each of the four educational levels. They made the anchor test and one of the two texts suited for their school level. (Students at the VWO level made the anchor test and one of the tests designed for HAVO. These students were included to obtain information on the discrimination of these tests at levels above the three operationalized levels.)

Results of phase 1

All pupils made the anchor test, so the calibration of items was done on a common scale. Since five anchor-items had fixed parameters, the scale was linked to the national sample. Beside the two texts, measures for (convergent) validation were administered: a vocabulary test, a verbal reasoning test (both GIT) and a questionnaire in which information was gathered about reading habits and school grades. These measures were helpful in selecting the most promising tests. IRT- and other analyses (summed estimated average information, number of malfunctioning distractors, factoranalyses on tetrachoric correlations, Mokken analysis) resulted in the selection of three texts: texts A and B at LBO/MAVO-level (12 items each) and text C at HAVO/VWO level (15 items).

Phase 2

The three texts A, B and C were administered to 1303 pupils of the major educational levels in three combinations: 380 AB, 542 AC and 381 BC. The same measures for validation were administered.

Results of phase 2

The item-parameters were calibrated with the same BILOG estimation procedure of Bayes' Modal Marginal A Posteriori with the same assumptions as mentioned before.

Analyses on all items (BILOG quasi-χ^2_{333}=811.5, p=.000) revealed that the 3-parameter-logistic-model will fit the data reasonably well, provided that a few non-fitting items are discarded. Further analyses (Mokken and LISREL unidimensionality checks, item- & ability parameter-invariance checks, IRT itembias and factoranalytic itembias analysis (Oort, 1990)) resulted in the selection of putative unbiased and efficient items, which maximally yielded about 20 points of information on a scale with standard normally distributed itemdifficulty.

Conclusions and Discussion

We succeeded in a reliable calibration of the items of the three texts, which form a putative valid wide range test of reading comprehension. However, for three reasons the accuracy of measurement is not as good as was intended.

Firstly, the estimated information that we achieved peaks rather sharply around the centre of the ability distribution and results in relatively little information around the

extremes of the ability distribution. Therefore, since the standard error of estimation of ability is the square root of the inverse of the information, the estimation of ability in the tails of the distribution will be less precise than was aimed at.
Secondly, around the center of the ability distribution, there *seems* to be enough information to estimate ability fairly reliable. However, the assumption of strict unidimensionality was never accomplished. What we achieved is 'relative unidimensionality' with respect to sex and level of education. Therefore the estimated maximum and average information will be an over-estimation of the true information contained in the test. Therefore the estimation of ability will be less precise than our testinformation suggests.
Thirdly, the estimated testinformation is based on *all remaining* items, which belong to the three texts. We administered just two texts to pupils, which took them about 50 minutes. An estimation of ability based on two texts will yield substantially less information and thus will be less accurate. To achieve the (overestimated) information by administering all three texts will require about 75 minutes.

References

Berkel, H. J. M. (1984). De diagnose van toetsvragen. [The diagnosis of testitems] Unpublished doctoral dissertation, Universiteit van Amsterdam, Amsterdam.

Groot, A. D. de, & Naerssen, R.F. van (1969). Studietoetsen. Construeren, afnemen, analyseren. [Tests. Construction, administration, analysis.] Den Haag : Mouton.

Kok, J., Krieken, R. v., & Luijten, A. J. M. (1986). Het construeren van open vragen. [The construction of open-ended questions] (Algemene publicatie nr. 45). Arnhem: Cito.

Kreeft, H. P. J. (1985). Een beschrijvingsschema voor tekstbegripvragen. [A desciptive scheme for reading comprehension questions] (Interne documentatie nr 152). Arnhem: Cito.

Mellenbergh, G. J. (1971) Studies in studietoetsen. [Studies on tests.] Unpublished doctoral dissertation, University of Amsterdam, Amsterdam.

Oort, F. J. (1991). Het gebruik van restrictieve factor analyse voor de constructie van een eendimensionale schaal volgens de Schenderstheorie. [Using RFA to construct a unidimensional scale in accordance with the Theory of Violators]. In: Hoogstraten, J., & Linden, W. J. van der (Eds.). ORD' 91: Methodologie. Amsterdam: SCO.

Oort, F. J. (in press). Using restricted factor analysis to detect item bias. Methodika.

Verhoeven, N. G., & Beuk, C. H. (1986). Het construeren van meerkeuzevragen. [The construction of multiple choice questions] (Algemene publicatie nr.29). Arnhem: Cito.

PSYCHOMETRIC ANALYSIS OF THE TURKISH GROUP ASSESSMENT OF LOGICAL THINKING (GALT) TEST

Giray Berberoulu, Middle East Technical University, Ankara, Turkey, & Frans J. Oort, University of Amsterdam, Amsterdam, The Netherlands

Assessment of logical thinking skills of students is one of the most important aspects of science teaching. As students are markedly different in the ability to grasp the concepts of science courses, logical thinking skills should be considered as one of the most prominent variables to be controlled and assessed within school curricula. Regarding the importance of this variable in education, Roadrangka, Yeany & Padilla (1983) developed the Group Assessment of Logical Thinking (GALT) test. Aksu, Berberoulu & Paykoç (1991) translated the GALT test into Turkish. Research indicates that both the original and the Turkish version of the GALT test are reliable and valid (Roadrangka et al., 1983; Aksu et al., 1991), and that the psychometric characteristics of two versions are highly consistent (Aksu et al., 1991).

These conclusions are based upon the results of the psychometric analyses that are common under classical test theory. It is purpose of the present paper to reanalyse the Turkish GALT test using the more modern methods of item response theory. Psychometric analyses will be carried out within the framework of the Theory of Violators (Oort, 1991). As a result the Turkish GALT test is shortened to become unidimensional and efficient.

Procedure

According to the Theory of Violators, it takes three steps to achieve unidimensionality and efficiency. Firstly, operationalize potential violators; secondly, detect biased items and remove them from the test; thirdly, detect and remove inefficient items. A potential violator is a variable with respect to which a test item may be biased. The Theory of Violators prescribes that, for a test to be unidimensional, item bias has to be investigated with respect to all relevant potential violators (Oort, 1991). As the GALT test will often be used to compare levels of cognitive development of different groups of students, students' sex and education level are very relevant and will serve as potential violators.

Subjects

Subjects are 1293 Turkish students from Ankara. 732 students are in the lower education level (junior high school) and 561 students are in the higher education level (high school). There are 686 males and 507 females.

Instrument
The GALT test consists of 21 items and measures six logical operations: conservation (4 items), proportional reasoning (6), controlling variables (4), probabilistic reasoning (2), correlational reasoning (2), and combinatorial reasoning (3). The first 18 items are in multiple choice format. To score, respondents have to select the correctanswer and justify their selection. The format of the last three items is open. All items are scored 0 (false) or 1 (correct).

Analysis
All analyses are done with the aid of Birnbaum's two–parameter logistic item response model and the computer program BILOG (Mislevy & Bock, 1986). Item bias is investigated with respect to one potential violator at a time. This means that the Birnbaum model is fitted to the data of four groups: 686 males, 507 females, 732 junior high school students, and 561 high school students. After re–scaling the parameter estimates, indices of sex bias and education level bias can be computed. Four different indices, based on the area between item characteristic curves are computed: AREA, DIFF (signed area), WAREA (weighted area), and WDIFF (weighted signed area).
After the removal of biased items from the test, the Birnbaum model is fitted again to the data of all students together. BILOG's item fit statistics and the maximum information values are used to select the most efficient items to construct a short version of the GALT test.
Further analyses concern the comparison of the different indices and the effect of shortening the test on testing for group differences.

Results
At the distribution of the area indices of item bias is unknown, item bias is evaluated by visually inspecting the coefficients (looking for outliers). In this way, possible bias is found in three items with respect to sex, in two items with respect to the level of education, and in one item with respect to both sex and education level. Four times (out of seven) the occurrent bias is against the lower ability groups (females, juniors). One item is biased against males. In two items the bias has no clear direction.
Not a single item is indicated as biased by all four area indices. In four instances bias is indicated by both the unweighted indices (AREA and DIFF), and in the other three instances bias is indicated by the weighted signed index (WDIFF). Rank correlations among the four indices show associations between AREA and DIFF (.61 and .89 for sex bias and grade bias respectively), WAREA and WDIFF (.77 and .88), and DIFF and WDIFF (.71 and .73).
To make sure that there will be no biased items in the shortened version, all six possibly biased items are removed from the test. On base of the item fit indices and maximum information values that result from fitting the Birnbaum model to the remaining 15 items, four more items are removed from the test, leaving a ten item test that may serve as a short form of the GALT test.
In order to compare the long and short forms of the GALT test, an ANOVA with sex and education level as independent factors is carried out two times: Once on the sum of all 21 GALT test items, and once on the sum of the ten items that constitute the short form. The ANOVA's indicate that shortening did not have a large effect.

Discussion

The analyses to detect bias show different results depending on what bias indices are used. The differences among the bias indices can be traced back to their nature. Differences between signed and unsigned indices may be due to nonuniform bias in test items. Differences between weighted and unweighted indices are due tothe fact that the largest differences between item characteristic curves are often found outside the range of the distribution of the person parameters.

The ten item short form of the GALT test seems to operate in much the same way as the full test. On the one hand this is comforting, but on the other hand it is discomforting that so much work selecting unbiased and efficient items has had so little effect. Maybe we have been too strict judging items as biased.

References

O, Berberoulu, G. & Paykoç, F. (1991). Can the GALT test be used in a different cultural setting? Research Report Middle East Technical University, Ankara, Turkey.

Mislevy, R.J. & Bock, R.D. (1986). *PC-BILOG.* Item analysis and test scoring with binary logistic models. Mooresville, IN: Scientific Software, Inc.

Oort, F.J. (1991). Theory of violators: Assessing unidimensionality of psychological measures. Paper presented at 7th European Meeting of the Psychometric Society, Trier, Germany.

Roadrangka, V., Yeany, R.H. & Padilla, M.J. (1983). The construction and validation of group assessment of logical thinking (GALT). Paper presented at the annual meeting of the National Association for Research in Science Teaching, Dallas, Texas.

ADJUSTING FOR MEASUREMENT UNRELIABILITY IN MULTILEVEL MODELS

M. Yang, G. Woodhouse, H. Goldstein, H. Pan, & J. Rasbash

Many measurements made in Educational research are subject to errors of measurement. This has long been recognised as an issue in achievement testing and there is a substantial literature concerned with methods of estimating the extent of such error variation. The present paper is concerned with the effects of measurement error on multilevel model estimates, and will start from the assumption that suitable estimates of measurement error variation are available. For a more detailed discussion of issues surrounding the estimation of such variation see Ecob and Goldstein (1983).
The simplest model takes the following form

$$x_i = t_i + e_i$$

where x_i, is the observed measurement on the *i - th* individual who has a true value for the measurement of t_i. The term e_i represents a random error of measurement. The reliability of X is defined, for a given population, as

$$R_X = \frac{var(T)}{var(X)}$$

When a variable containing measurement error is used as a predictor or explanatory variable in a linear model then certain difficulties arise. In the case of a simple linear regression model

$$y_j = \alpha + \beta t_j + e_j$$

the variable T cannot be observed and we cannot directly obtain consistent estimates of in the usual way. We can, however, estimate ' 'in the model

$$y_j = \alpha' + \beta' x_j + e'$$

and it is known that a consistent estimator of is given by $1/R_x$ where R_x is known or a consistent estimator is available.
If R_x is substantially less than 1.0 then failure to make this adjustment can lead to misleading inferences. A comprehensive account of dealing with measurement errors in single level models is given in Fuller (1987).
Turning now to the multilevel case, we encounter similar problems when explanatory or response variables have errors of measurement. In addition, the estimates of the random parameters, that is the variances and covariances at each level are also affected. Goldstein (1986) derives procedures for taking into account measurement errors when they occur in variables measured at level 1. This might occur, for example, in a 2-1evel model of students grouped within schools where a prior achievement measure is used as a covariate. The level 1 procedures have now been extended to handle measurement errors in variables measured at higher levels.
Thus, for example, suppose we wished to use the 'contextual' variable which is the average of the prior achievements of all the students in a class, as in the following model

$$y_{ij} = \alpha + \beta x_{ij} + \gamma \bar{x}_{.j} + u_j + e_{ij}$$

where $x_{.j}$ is the class mean prior achievement score and u_j, e_{ij} respectively are the level 2 and level 1 random residual terms. If only 50% of the students in each class had been sampled, then we would have only an *estimate* of $x_{.j}$ and the between-class variance of this estimate would be twice that of the variance of the true value, that is it would have a reliability of 0.5.
The need to adjust for level 2 measurement errors as well as those for level 1 variables occurs in many fields of application. In household surveys, where we wish to model the hierarchical population structure, we typically will sample only a small percentage of households in an area. If we wish to use area level variables, based upon aggregating the sample household characteristics, then these aggregate level variables will tend to have low reliabilities.
The paper will present basic results and formulae and show how these can be used to make necessary adjustments. It will illustrate the effect of taking account of level 1 and level 2 measurement errors using a longitudinal data set dealing with the educational attainments of primary school children and will demonstrate that failure to take account of measurement error can lead to misleading interpretations.

References

Fuller, W. A. (1987). Measurement Error Models. New York: Wiley.
Ecob, R., & Goldstein, H. (1983). Instrumental variable methods for the estimation of test score reliability. J. Educ. Statist., 8, 223-241.
Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalised least squares. Biometrika, 73, 43-56.

APPLICATIONS OF MULTIPLE-GROUP IRT MODELS

Michele F. Zimowski, National Opinion Research Centre (NORC), & R. Darrell Bock, University of Chicago, USA

Abstract

Multiple-group IRT models have many uses in educational measurement. They include 1) anchor-item equating of test forms administered to different populations, 2) vertical scaling over grade or age groups, 3) detection and estimation of differential item functioning (DIF), 4) estimation of item parameter drift, 5) two-stage testing, and 6) calibration of sections of test forms that have been self-chosen by the examinee. All the familiar IRT models for dichotomous and polytomous item responses can be calibrated for these applications by the maximum marginal likelihood method if separate latent distributions are assumed for each group. Concomitantly, latent densities for the groups can be estimated parametrically or nonparametrically on a common scale. Scores of the examinees may then be estimated on that scale without regard to the examinee's group membership if maximum likelihood is employed, or by Bayes estimation if group membership is observed.

The paper presents and discusses a number of applications of multigroup logistic IRT models. Both simulated and real data are analyzed. This summary deals with one of these examples that has been constructed from a sample of actual test data in such a way as to simulate the first and second stage of a two-stage test administration.

The two-stage procedure is a form of adaptive testing suitable for paper-and-pencil administration (see Lord, 1980). The examinees respond to a short first-stage test at an earlier time and to a second-stage test at a later time. The level of difficulty of the second-stage test is adapted to each examinee's proficiency-group assignment as determined from the first-stage score. The forms of the second-stage test range from easy to difficult, and adjacent forms are connected by common items for purposes of IRT scaling.

From point of view of item calibration, two-stage testing is formally identical to test equating by the anchor-item method when the forms are administered to persons from different populations. It is also the same as vertical equating of tests for different age groups when the tests are connected by common items. Although it is possible to analyze data from these types of applications by fixed-effect maximum likelihood estimation, as in Lord (1980), the marginal maximum likelihood approach employed in this study is more informative, more stable numerically, and

784

has better-defined asymptotic properties. The following example illustrates these types of important multiple-group IRT applications.

Data

Data for this example were created from a large sample of subjects responding to a 100 word spelling test. The words of the test were presented by a paced tape recording, and the subjects wrote the words on a numbered answer sheet. The answers were scored as correct or incorrect, and incorrect responses were coded by type of error. In the present analysis, we use only the right-wrong scoring. Because these are constructed response items and guessing is not possible, the two-parameter logistic model is suitable.

Items 9, 10, 17, 54, 92, and 95 of the spelling test were selected to represent a six-item pretest. Examinees with scores 0 , 1, and 2 were assigned to second-stage form 1 (easy), those with scores 3 and 4 to form 2 (intermediate), and those with scores 4 and 5 to form 3 (difficult). The numbers of subjects assigned to the second-stage forms by the pretest were 128, 396, and 136, for Forms 1, 2, and 3, respectively.

Eight items were assigned on the basis of their sample *p*-values to each of the three, second-stage forms as follows:

- Easy 1, 34, 36, 39, 59, 64, 85, 87
- Intermediate 15, 27, 33, 34, 49, 48, 64, 79
- Hard 8, 15, 23, 25, 28, 47, 72, 79

Common items 34 and 64 connect forms 1 and 2; items 15 and 79 connect forms 2 and 3.

Method

When the second-stage items are calibrated by maximum marginal likelihood, separate latent distributions must be estimated for the groups of examinees defined by the assignments to second-stage forms. This type of calibration has been implemented in the BIMAIN 2.0 program of Zimowski, Muraki, Mislevy and Bock (1991). The program estimates the mean, standard deviations, and point histogram of the latent distribution simultaneously with the item parameters. The indeterminacy of the location and scale of the latent variable is resolved either by setting the mean and standard deviation of the distribution for one of the groups to 0 and 1, respectively, or, in the case of two-stage testing where the subjects originally came from one distribution, by setting the mean and standard deviation to these values in the combined estimated distributions of the groups weighted by their sample sizes. The program can perform these analysis for as many subtests as the investigator wishes to compose from the original item-response data.

In the scoring phase of the program, scale scores for the examinees may be estimated by maximum likelihood, in which case the group membership of the subject is irrelevant. Alternatively, Bayes estimation may be used with the latent distributions from the calibration phase serving as priors according to the group membership of the subject; in this way, the information contained in the classification of the subjects by the first-stage test is carried into the second-stage scores. In either case, the location and scale of the estimated scores can be adjusted to any arbitrary mean and standard deviation in the general distribution for all subjects.

As a check on the validity of the solution, we have calibrated the data first as two-stage test results, then, using all twenty items, as a conventional test administered to one group of examinees. The similarity of the item parameter estimates and the latent distributions under these two conditions shows the extent to which the one- and two-stage analyses are consistent.

Results and Discussion

The classical item *p*-values in the two-stage data are more uniform and intermediate in range than in the one-group results. This is the main advantage of two-stage testing: it assigns the items to the second-stage groups so that the probability of correct response among those examinees who are presented the items will be closer to 50 percent, and thus more informative.

Despite the smaller amount of data supplied by the two-stage test compared to the one-group test (test length 8 versus 24, and smaller sample sizes per item), the item parameter estimates from the two analyses are generally in good agreement. The effect of the relative amounts of data is apparent in the larger standard errors of the two-stage results. As would be expected because of the restricted range of ability within the second-stage groups, agreement of the thresholds is better than that of the slopes.

In an actual application of two-stage testing, many more examinees per second stage group would be available for calibrating the instrument and the number of items per form would be perhaps 15 with at least 3 shared items between adjacent forms. In that case, the concordance between two-stage and one-stage results should be more than adequate to justify the reduction in testing time afforded by the two-stage testing procedure.

The estimated item parameters in the two-stage analysis were scaled so that the combined latent distributions of the groups, weighted by the sample sizes, have zero mean and unit standard deviation. In this scale, they can be compared with the estimates for the one-group analysis with a similarly scaled latent distribution. Although the points at which the latent densities were evaluated are slightly different in the two-stage and one-group analyses, it is apparent that the form of the combined two-stage distribution is highly similar to that of the one-group distribution. The differences are due almost entirely to the more diffuse likelihoods for the ability values in the responses to 8 items versus 24 items. The results illustrate the consistency of the maximum marginal likelihood method in estimating the form of the latent distribution along with the item parameters, even when the tests differ considerably in length.

References

Lord, F. M. (1980). Applications of item response theory to practical testing problems. Hillsdale, NJ: Erlbaum.

Zimowski, M.F., Muraki, E., Mislevy, R.J., & Bock, R. D. (1991). BIMAIN 2.0. Test maintenance: item calibration, test equating, differential item functioning, item parameter drift, and variant item statistics. Chicago: Scientific Software, Inc.

ON THE EXISTENCE OR OTHERWISE OF COMPOSITIONAL EFFECTS IN MULTILEVEL MODELLING OF SCHOOL EFFECTIVENESS

Dougal Hutchison, & Ian Schagen, National Foundation for Educational Research in England and Wales, United Kingdom

Recently there have been a number of projects which have aimed to look at the progress or 'value added' in schools in England and in Scotland. These have used multilevel modelling to compare actual and predicted attainment, using a variety of predictors, generally either some measure of home background such as socioeconomic group, or parental education, or, preferably, some measure of prior attainment, such as an attainment test at or shortly before entry to school. One aspect of interest has been the question of compositional effects. In this we follow the terminology of Bryk & Raudenbush (1988). These may be defined as the extent to which collective properties of a pupil body have an effect on pupil attainment, over and above the effect of individual pupil characteristics. A number of writers have considered the effect of such ecological variables. Among those who concluded that it was an important effect were Rutter et al (1979), Willms (1985). Smith & Tomlinson (1989) did not find a statistically significant effect, but argued that this apparent failure might be due to the relatively small size of the sample, while Bondi (1991) found an effect with one group of primary school children in Scotland, but not with another.

A study by Gray et al (1990) emphasised that the question of ecological effects has to be treated with caution. Using data for six local education authorities in England, they found that while some of their authorities apparently showed substantial ecological effects on progress, others did not. Significantly they found that the effects of introducing ecological effects into the model were virtually undetectable when there was already a good measure of prior attainment. While they stop short of actually saying that ecological effects do not exist, they conclude that 'the evidence for strong "compositional" or "contextual" effects is relatively weak'. Their implication is that many of the apparent ecological effects arise because the measure of intake quality is poor, and the apparent ecological effects are simply unexplained individual level differences being 'mopped up' by the aggregated variable.

These results are somewhat confusing. Does this mean that we can never be sure that there in fact is an ecological effect, that any apparent effect might simply be be an artefact arising because one can never measure all the precursors of success in the future? On the one hand, the arguments of Gray and his co-workers suggest rather strongly that the evidence in favour of ecological effects has to be

treated with some caution, to say the least. On the other hand, common sense, and one's own experience, tend to be convincing advocates of a strong effect of the attainment and attitudes of one's classmates. With this background, we report some results from a study of school effectiveness which have, in our opinion, an important bearing on this question.

The Data

The NFER, in conjunction with an LEA in the South of England, has recently carried out a series of studies of school effectiveness in teaching reading, as a by-product of an authority-wide programme of testing in reading. In this paper, we describe and compare results from two cohorts of some 2500 children in some 50 schools. A wide variety of data was collected, but what is important here that reading scores were collected at two-year intervals for the same children, at 6 and again at 8 in Cohort 1; and at 8 and again at 10 in cohort 2. Of those for whom a record was returned, over 98 per cent supplied a test score result for 1990, and of these about 85 per cent supplied a test score from 1988.

To investigate compositional effects, pupil attainment in 1988 is recalculated as a deviation from the school mean, and the latter is included as an additional school-level variable. If there is a compositional effect, then one would expect that it would manifest itself by the school-level regression coefficient being larger than that for pupil-level attainment.

The Table compares the results for Cohort 1 (Column a), and Cohort 2 (Column 2). The dependent variable is reading score in 1990, age-standardised, and converted to a mean of 100 and a standard deviation of 15. The table shows us that there is, as might be expected, a strong connection between 1990 reading score, and score on another test in the same series two years earlier. The within-school pupil-level regression coefficient for 1988 Score is of the order of .8. We see also that while there is no statistically significant difference and thus no compositional effect is suggested for Cohort 1. On the other hand, the coefficient for mean class score in Cohort 2 is 1.06, significantly larger than the individual coefficient, indicating the existence of a compositional effect.

Table

Outcome Variable	Standardised score in 1990			
	Cohort 1		Cohort 2	
Fixed Effects				
Pupil-level				
1988 Score	0.81	(.01)	.83	(.01)
Sex (boys)	1.3	(.3)	.74	(0.3)
Class-level				
1988 Mean class score	.81	(.05)	1.06	(.05)
Random Effects				
Pupil-level	56.4	(1.7)	43.8	(1.4)
School-level	3.8	(1.0)	2.4	(0.7)

This is a particularly interesting finding. An apparent compositional effect in one cohort in isolation could be explained away as an unavoidable statistical artefact. However, the lack of evidence for a compositional effect in the younger cohort is

particularly significant: it is not inevitable that an artefactual compositional effect will appear.

Discussion

This paper indicates strongly that compositional effects in school effectiveness do exist, as would be predicted by personal experience. The main reason that we have been able virtually to rule out the existence of artefactual effects is that we have a particularly strong predictor. Gray recommends a correlation of at least 0.7 between predictor and outcome, whereas the prediction in both our cohorts has been over .8.

Is there any explanation of the discrepancy between the two cohorts. It is of interest to note that while, as noted above, Bondi's (1991) suggestion that compositional effects are more notable in older children has been confirmed by our study. While the reasons must remain a field for the exploration of child psychologists, a tentative suggestion lies in the development of peer-group influences with age.

References

Bondi, E (1991) 'Attainment at Primary Schools: an analysis of variation between schools'. British Journal of Educational Research, 203-218.

Gray, J; Jesson, D & Sime, N (1990) 'Estimating differences in the examination performance of secondary schools in six LEAs: a multilevel approach to school effectiveness' Oxford Review of Education, 16,137-158.

Mortimore, P; Sammons, P; Stoll, L; Lewis, D & Ecob, R (1988) School Matters (London, Open Books).

Nuttall, D; Goldstein, H; Prosser, R & Rasbash, J (1989) 'Differential school effectiveness' International Journal of Education Research, 13, 769-776.

Bryk, A & Raudenbush, S (1988) 'Methodology for cross-level organisation research' Research in Organisational Behaviour.

Rutter, M; Maughan, B; Mortimore, P & Ouston, J (1979) Fifteen Thousand Hours: secondary schools and their effects on children London: Open Books.

Smith, D & Tomlinson, S (1989) The School Effect: a study of multi-racial comprehensives. London: Policy Studies Institute.

Willms, D (1985) 'The Balance Thesis: contextual effects of ability on pupils' O-grade examination results', Oxford Review of Education, 11, 33-41.

AN APPLICATION OF A MULTILEVEL MODEL IN THE ANALYSIS OF PANEL DATA

Marcel W. Vooijs, Cees M. Koolstra, Centre for Child and Media Studie, & Leo J.Th. van der Kamp, Department of Psychology, Leiden University, Leiden, The Netherlands

Abstract

Multilevel analysis meets several drawbacks of traditional techniques for the analysis of quantitative panel data. In this paper the results are reported of the application of a multilevel model for the analysis of panel data in a study of the effect of watching TV and reading in free time on reading proficiency. This panel study comprises two age cohorts from sixteen primary schools. Data have been collected on three occasions with a time lag of one year between two consecutive occasions. Based on the results of this analysis, the advantages of a multilevel approach for the analysis of panel data are discussed.

A method for the design of longitudinal research is a panel study. In a panel study each individual is observed on more than one occasion. The observation may be a quantitative measurement, or a state, such as presence or absence of certain characteristics (e.g. symptoms). Usually a group of subjects is studied, and quantitative data are collected on one or more occasions with the aim to study changes in relevant variables over time. Several methods for the analysis of quantitative panel data have been proposed: ANOVA, MANOVA, multiple regression analysis, and LISREL. These methods are all based on the general linear model. There are a number of drawbacks, however, in using them for the analysis of panel data.

Firstly, the techniques mentioned are based on a static model of analysis. The time of measurement is conceived of as a discrete variable, and the analysis takes merely the scores on the discrete time points into consideration. The dynamic character of intra-individual changes is not considered explicitly, i.e. there is no intrinsic interest in age- or time-related changes in the variables. Recently a number of approaches have been proposed in which the time variable has been explicitly taken into account or modelled (e.g. Bock, 1976; Guire & Kowalski, 1979; Rogosa & Willett, 1985; Bryk & Raudenbush, 1987, 1989; Raudenbush & Bryk, 1988). These approaches have advantages for describing and explaining temporal changes in the scores on the variables under study, stressing the process of change rather than the end result of it.

Secondly, traditional techniques neglect the so-called hierarchical or multilevel character of the data. In educational panel research e.g. choosing a sample often

amounts to selecting schools, and after the schools have been selected, classes and students within classes are selected of which data are collected. Clearly, the sample at hand is hierarchically organized: the students are nested within the classes, and the classes are nested within the schools. Hence individual scores within schools are not independent of each other, and the distribution of scores among schools may differ. In the traditional methods of analysis it is assumed that the observations are independent of the specific schools under investigation. Consequently the effects are estimated less precisely.

In summary, the two principal motivations to design a panel study longitudinally and analyze the data correspondingly are: (1) to examine the individual's changing response over time paying explicitly atttention to the process character of the study, (2) to increase the precision of treatment contrasts by eliminating interindividual variabililty from the comparisons of interest (see e.g. Cook & Ware, 1983).

Multilevel analysis meets the drawbacks of the traditional techniques of analysis mentioned. Multilevel models combine a within-subject model, characterizing individual growth, with a between-subject model. Several growth functions have been proposed for modelling the individual growth trajectories. Bock (1976) chooses the logistic funtion to describe individual growth; Rogosa and associates (Rogosa et al., 1982; Rogosa & Willett, 1985) propose polynomial models for characterizing individual growth; Thissen and Bock (1990) prefer exponential functions. In a multilevel model approach growth is specified as a random variable, and simultaneously the variables affecting growth (e.g. differences between schools) are specified. Taking into account multilevel effects in the statistical analysis of the data increases the precision of estimating the effects of interest (De Leeuw & Kreft, 1986; Goldstein, 1987, 1989; Raudenbush & Bryk, 1986, 1988).

In the present paper the results will be reported of the application of a multilevel model for the analysis of panel data in a study of the effect of watching TV and reading in free time on reading proficiency. This panel study comprises two age cohorts from sixteen primary schools. Data have been collected on three occasions with a time lag of one year between two consecutive occasions. At the first occasion, the younger cohort consisted of the children of group 4 of primary school, the elder cohort of the children of group 6. On each occasion the frequency of watching TV and reading during a certain period has been measured, as well as reading proficiency.

In the specification of the multilevel model utilized, three questions were preconceived: (1) how does reading proficiency develop over time, and are there individual differences? (2) are differences in the development of reading proficiency related to differences between schools, and to differences between cohorts? (3) to what extent are watching TV and reading in free time (besides sex) influencing the development of reading proficiency? Based on the results of this analysis, the advantages of a multilevel approach over traditional techniques for the analysis of panel data are discussed.

References

Bock, R.D. (1976). Basic issues in the measurement of change. In D.N.M. de Gruijter & L.J.Th. van der Kamp (Eds.), Advances in psychological and educational measurement. (pp. 75-96). London: Wiley.

Cook, N.R., & Ware, J.H. (1983). Design and analysis methods for longitudinal research. Annual Review of Public Health, 4, 1-23.

Bryk, A.S., & Raudenbush, S.W. (1987). Application of hierarchical linear models to assessing change. Psychological Bulletin, 101, 147-158.

De Leeuw, J., & Kreft, I. (1986). Random coefficient models for multilevel analysis. Journal of Educational Statistics, 11, 57-85.

Goldstein, H. (1989). Models for multilevel response variables with an application to growth curves. In R.D. Bock (Ed.), Multilevel analysis of educational data (pp. 107-125). San Diego: Academic Press.

Guire, K.E., & Kowalski, Ch.J. (1979). Mathematical description and representation of developmental change functions on the intra- and interindividual levels. In J.R. Nesselroade & P.B. Baltes (Eds.), Longitudinal research in the study of behavior and development (pp. 89-110). Nerw York: Academic Press.

Raudenbush, S.W., & Bryk, A.S. (1986). A hierarchical model for studying school effects. Sociology of Education, 59, 1-17.

Raudenbush, S.W., & Bryk, A.S. (1988). Methodological advances in analyzing the effects of schools and classrooms on student learning. In E.Z. Rothkopf (Ed.), Review of Research in Education (vol. 15), (pp. 424-475). Washington, D.C.: American Educational Research Association.

Rogosa, D.R., & Wilett, J.B. (1985). Understanding correlates of change by modeling individual differenes in growth. Psychometrika, 50, 203-228.

Rogosa, D.R., Brand, D., & Zimouski, M. (1982). A growth curve approach to the measurement of change. Psychological Bulletin, 90, 726-748.

Thissen, D., & Bock, R.D. (1990). Linear and nonlinear curve fitting. In A. von Eye (Ed.), Statistical methods in longitudinal research. Boston: Academic Press.

PROBLEMS IN THE EVALUATION OF THE ASSUMPTIONS OF THE MODEL USING BOX-PLOTS.

A.M. Lopez, & M. Morales, Department of Developmental and Educational Psychology, Basic Psychology and Methodology, University of Seville, Spain

Summary

The box plot is a technique developed by Tukey (1977) to explore the statistical characteristics of data such a symmetry, normality, etc. Its main advantage is that it allows us to summarize the distribution characteristics using robust indices (median and the fourths), while allowing us at the same time, to examine with great care the possible values of outliers.

For a symmetric distribution, the median equals the mean, so the population median of the standard Gaussian distribution is 0. The population fourths are -0.6745 and 0.6745, so the population fourth-spread is 1.349, or about 4/3. Thus 3/2 times the fourth-spread is 2.0235 (about 2).The population outlier cutoffs are between -2.698 and 2.698 (about 2 2/3), and they contain 99.3% of the distribution.

Emerson and Strenio (1983) made a simulation study in order to compare population values of the median, fourths, and outlier cutoffs for several familiar distributions (uniform, standard Gaussian, t-distributions and chi-squared distributions).

They found that for large samples the value of the median coincided in all the symmetrical distributions studied, while it was different in the asymmetrical ones. Thus, we found that the limit for the outlier cutoff was even lower than the lowest score in the assymetrical distributions. The authors concluded that, for large samples, the assymetrical distributions may be detected in the box-plot through the situation of the median and the location of the outliers in just one of the distribution tails.

Using these results as our base, we propose using the box-plot to detect the normality of the empirical distributions. To do this, we started by recognising that in educational research it is not always possible to work with large samples, it being normal to carry out studies with samples sizes n _ 100. In this way, we considered that the empirical distributions obtained in educational research does not always respond to the standard distribution model N (0,1). As a result, the aim of our research was to study the normality of the distributions using box-plots in distributions of sample sizes n _ 100 and with different values in the central tendency parameters and dispersion.

Method

To test the suitability of the box-plot in the description of the sample characteristics of one variable (central tendency, dispersion and form), we carried out an experiment of Monte Carlo.

Using normal distributions with different averages and deviation types (Averages equal to 0, 5, 10, 20, 40 and 80; SD = 1, 2, 3, 4, 5, 6 and 7) we generated samples of two sizes (50 and 100) to a total of 42 combinations. All the combinations were analysed building a box-plot, a histogram and submitting them to the Kolmogorov-Smirnov normality test. For the different analyses we used the statistical packet SPSS/PC+ v. 3.1.

Results

We found that the inference of normality of the samples could be established in the 42 combinations studied using the histogram and the significance test. The latter with probabilities for the null hypothesis over 0.7.

However, the box-plots of some of the samples (average equal to 0 and standard deviation (S.D.) = 3; average equal to 0 and SD = 7, average equal to 10 and SD = 1; average equal to 10 and SD = 3; average equal to 20 and SD = 1) show clear violations of the assumption of normality of the samples (assymetries, tail elimination, coincidence of the median with the first of third quarter) that contradicted the results obtained with other procedures.

The different conclusions which we can draw, depending on the procedure we use to infer normality, led us to analyse the possible problems of the algorithm of construction of the box-plot implemented in the statistical packet SPSS/PC+, V. 3.1, as well as the search for procedures for the consruction of the box-plot that solve the problems detected in the present research.

References

Tukey, J.W. (1977). Exploratory data analysis. N. York: Wiley.

Emerson, J.D. & Strenio, J. (1983). Box-plots and batch comparison. In D.C. Hoaglin, F. Mosteller & J.W. Tukey (eds.). Understanding Robust and exploratory data analysis. N. York: Wiley.

USES OF QUALITATIVE/QUANTITATIVE TERMS IN EDUCATIONAL RESEARCH.

M. Morales, & A. Moreno, Department of Developmental and Educational Psychology, Basic Psychology and Methodology, Universidad de Sevilla, Spain

Summary

In recent years a vigorous debate has emerged among researchers in Humanities concerning the problems presented by quantitative methodology and the suitability of substituting or complementing this with a qualitative methodology. (Taylor & Bogdan, 1975; Patton, 1984; Guba, 1982; Guba & Lincoln, 1982; Anguera, 1985; Moreno, 1987, etc.).

It cannot be said, however, that this debate has been concluded or that there is even a certain level of consensus, given that a series of problems exists (ambiguity and polysemy in the use of terminology, lack of systematization of the various meanings), that prevents the establishing of a clear conceptualization of both methodologies.

As a result, discrepancies arise among authors relating to subjects as diverse as that of the complementariety or not of both methodologies, the meaning given to each one of the terms, the suitability or not of using triangulation, etc.

Thus, it is difficult to accept the proposal put forward by Smith and Heshusius (1986) of considering the debate between the two methodologies closed, as for many authors it is still not known what the terms mean and nor is there any consensus about them. Furthermore, it may be stated that the above-mentioned problems are responsible for the level of confusion that exists within the field of study (Smith, 1983; Jacobs, 1988), it being necessary to put forward solutions for each of them if one wishes to achieve a correct conceptualization of these methodologies.

With this aim in view, we have carried out an exploratory content anlysis of the characteristics most frequently used by authors to delimit the terms C and Q. The results obtained support the hypothesis formulated by Reichardt and Cook (1986) that the terms C and Q are used as much on a paradigmatic as a methodological level. The only difference between one author and another seems to lie in the particular characteristics selected, it being found that there are two specific forms of delimiting each one of the said levels.

However, this affirmation must be qualified for several reasons: the first is that, although the paradigmatic and methodological levels seem sufficiently contrasted, there is no evidence to suggest that they always appear with the characteristics

proposed by Reichardt and Cook (1986). In many cases, the terms are mentioned with only some of the characteristics put forward by the above-mentioned authors.
The second objection we have found in this study to Reichardt and Cook's proposal is that authors do not always delimit the terms C and Q with the characteristics proposed by them. More specifically, in this study we observed that the term Q was sometimes delimited by characteristics of the term quantitative (data analysis through Statistics, supporters of exact measurements).
The third is that the variance associated to the five factors found in the empirical study, while being acceptable, does not allow them to be considered as definitive or conclusive. From our point of view, it is not very probable that the rest of the variance not associated to the factors is purely accidental. It is possible that part of it may be related to other classification criteria.
To sum up then, we consider that the empirical study illustrates that Reichardt and Cook's proposal (1986), although representing a significant advance in the field, does not go far enough, given that it does not gather all the possible uses of that authors attribute to the terms C and Q. In our opinion, this may be considerd as proof of the need to look for new criteria and/or reformulate the alredy existing ones, in such a way that we achieve a complete enumeration of the uses of the terms in question.
The problems that Reichardt and Cook's proposal presents may be overcome, in our opinion, if the different uses of the terms C and Q are considered to be the result of the following dimensions: a) the paradigmatic notion that the researcher uses, and b) the attitude each author adopts faced with the set of problems.
The first of them may be considered as a re-elaboration of Reichardt and Cook's proposal with the end of saving the previously mentioned difficulties. The second criteria, on the other hand, has not yet been put forward, but it is necessary in order to understand those uses in which the terms C and Q are delimited with characteristics which do not correspond to them according to Reichardt and Cook's proposal (1986) or to understand integrating postures such as those of Weber and his supporters.

References

Anguera, M.T. (1985). Posibilidades de la metodología cuantitativa vs. cualitativa. Revista de Investigación Educativa,3 (6), 127-144.

Guba, E. (1982). Criterios de credibilidad en la investigación naturalista. En Gimeno, J. y Pérez, A. (comps.). La enseñanza: su teoría y su práctica. Madrid: Taurus.

Guba, E. & Lincoln, Y.S. (1982). Effective Evaluation. San Francisco: Jossey-Bass.

Jacobs, E. (1988). Clariffying qualitative research: A focus of traditions. Educational Researcher. (Jn-February), 16-24.

Moreno, R. (1987). Hacia una superación de los paradigmas cuantitativo y cualitativo. Comunicación presentada al II Congreso de Evaluación Psicológica. Madrid.

Patton, M. (1984). Qualitative evaluation methods. Beverly Hills: Sage.

Reichardt, CH.S. & Cook, D. (1986). Hacia una superación del enfrentamiento entre los métodos cualitativos y los cuantitativos. En T.D. Cook y Ch.S. Reichardt (eds.). Métodos cualitativos y cuantitativos en investigación evaluativa. Madrid: Morata.

Smith J. K. (1983). Quantitative versus interpretative: The problem of conducting social inquiry. En E. R. House (ed.). Philosophy of evaluation. San Francisco: Jossey Bass.

Smith, J. K. y Heshusius, L. (1986). Closing down the conversation: The end of the quantitative- qualitative debate among educational inquirers. Educational Researcher, 15, (1), 4-12.

Taylor, S. & Bogdan, R. (1984). Introduction to qualitative research methods. New York: Wiley.

NEW MEANS OF FEEDBACK EVALUATION IN RESEARCH ON EDUCATION: RESEARCH METHODOLOGY AND EVALUATION RESEARCH

K.R.G. Ulbricht, University of the Federal Armed Forces, Faculty of Education, Munich, Germany

Abstract

A newly developed method and technical device is described which allows to collect feedback information from the students within the classroom and which permits to evaluate these data immediately. Now, equipped with this device, the classroom teacher is able to react without any delay. Furthermore, the lecture can easily be documented and analyzed later in more detail.

Objective data (solutions to given problems) have been correlated within the framework of our study to subjective data (ratings of the students), and they showed significant correlation resp. contingency coefficients. Another application is the measurement of special parameters of research about the classroom - looked upon as time-series - e.g. the "acceptance" or the "classroom tension" (Freytag-Loringhofen 1973) of "Didactical Contacts" within the classroom (Winnefeld 1973).

Introduction

One of the predominant problems of evaluation of the feedback of the students within the classroom is the lack of tools which enable the teacher to make a better and quicker use of this most important information.

The collection as well as the evaluation of the feedback data that stem from the students have been very time-consuming in the past. Moreover, a large amount of time was necessary to evaluate the classroom data which have been collected by the well-known paper-and-pencil tests resp. examinations, questionnaires, ratings of observers, written or oral reports etc.

It is desirable to use a method which allows quicker reactions of the teacher as well as more penetrating methods of research on teaching and learning.

Method

The method in order to reach these goals culminated in the development of a special technical device. Its most important components are:

- A central unit ("reactoscope") and a personal computer which is equipped with special software, which are both to be operated by the teacher resp. by his assistant.

- Several (e.g. 64) input devices, which are hand-held by the students. Every student makes her or his choice within the range of the scale at her or his input device, that allows two functions, namely:
- a continuous rating scale from -3 to +3 or
- a discrete choice from 1 to 7 (natural numbers).

The creations of Hedges & MacDougall (1964),Freytag-Loringhofen (1973), and Vehrs (1986) form some sort of forerunners of our device.

Empirical Studies and Results

The above-mentioned method and device was first applied at our University when a new repetitive course on "Statistics" was planned within the Faculty of Education. We decided to evaluate the pre-knowledge of the students first. Therefore, Multiple Choice Items were constructed with relation to the educational objectives, e.g. level 1.3 from "knowledge", level 2.2 from "comprehension", and 6.2 from "evaluation" (comp. Bloom et al 1971).

One and only one of the possible choices was the correct one, and the distractors referred to misconceptions ("negative educational objectives") which appear, in our experience, often.

The above-described device was applied at several experiments and at a participant observation within the classroom.

Experiment 1 (e.g.)

Every MC item consisted of four choices, wherefrom the fourth, "I do not know", was identical for all the 12 items. This study was anonymous and did not form part of any examination.

After having answered a MC item, the students were asked to give a rating of the item difficulty using a continuous rating scale (-3 = most difficult; 0 = indifferent i.e. not easy but also not difficult); +3 = very easy).

The last step, hypotheses testing, led to the following propositions:

- The subjective rating is a valid measure of the objective item difficulty.
- The number of given answers on "I do not know" is a valid measure of both the item difficulty and the rating of the difficulty.

The design and the outcomes of this and some other experiments will be discussed and illustrated by several overhead projections (photos, tables, and figures).

Participant Observation of a classroom lecture

Wilhelm Wundt, the creator of the psychological experiment and laboratory in Germany, has analyzed the system of human emotions thoroughly. He stated that "tension" and her annihilation plays an important role at emotions as well as at communication (Wundt 1896/1911, 1900). Freytag-Loringhofen transferred this approach to the classroom. His technical means were audio tapes and his "kymographion", i.e. a device, which allows to fix marks on a uniformly rotating paper reel.

Our reactoscope is a new means of doing research on descriptors of the classroom.

Discussion

Dr. Gall, who applied an identical device, evaluated the validity of the rating scale of the hand-held feedback devices. His result was the exact linearity when using the full range of the scale (Gall 1986).

Moreover, this researcher evaluated educational media (video and audio tapes) under the condition of time-dependency of the ratings of his students (Gall 1986, 1990). He (1988) as well as Dr. Vehrs (1986) who used a device which was comparable (it allowed only for one input) added time-dependent studies on the psychology of the arts.

Another important result from Vehrs was that the correlation of the ratings when using his mechanical input device was parallel in a significant way to the ratings when making the choice on a traditional questionnaire.

Furthermore, the results of other researchers, who applied kinds of forerunners of our device (Hedges & MacDougall 1964, Freytag-Loringhofen 1973, Vehrs 1986) will be summarized and discussed.

To summerize the discussion of the empirical studies using a mechanical and hand-held input device: The field of application in not exhausted by now, and the outcomes seem to possess a large amount of validity and reliability.

Conclusion

The conclusion of our studies applying the new method and device aimed at getting feedback information from the students immediately, as well as the studies of collegues using an identical or a similar device, could be that there are three main fields of application, namely:

- The field of education:
  The formative and summative evaluation of student lear- ning, esp. in the areas of knowledge, comprehension, and evaluation of the cognitive domain as well as in the affective domain of educational objectives (Bloom et al. 1971).
- The field of research on teaching and learning:
  The effect of different didactical approaches, and of interventions.
- The field of economical considerations, with relation to the evaluation of educational aids, material, and of the quality control aimed at optimizing it.

References

Bloom, B.S., Hastings, T.S. & Madaus, G.F. (Eds.) (1971). Handbook on formative and summative evaluation of student learning. New York.

Freytag-Loringhofen, W-D. von,. Zur Unterrichtsdynamik. In: Winnefeld, ibid., p. 81-105

Gall, S. (1986). Rükmeldung im Unterricht. Zur Konzeption eines Gerätes. In: Medien und Technik, Series Ingenieur- Pädagogik, ed. by A. Melezinek, Vol. 22, Alsbach, p. 138-141

Gall, S. (1988) Bewertung audio-visueller Medien. Diskussion von Test-Ergebnissen. In: Technik lehren - Technik lernen, Series Ingenieurpdagogik, ed. by A. Melezinek, Vol. 24, Alsbach, p. 135-140

800

Gall, S.(1990) Zeitreihenanalyse zur Akzeptanz von Unterrichtsmedien. In: Ingenieurausbildung 2000/Engineering Education 2000. Proceedings Internat. Symposium Ingenieurpädagogik 90/ Frontiers in Engineering Education 90, 20th Annual Conference, ed. by Haug/Melezinek/Schutz. Alsbach , p. 149-151

Hedges, W. D. & MacDougall, M.A., (1964) Recording Student Perceptions of Teacher Behavior By Means of a Student Response Monitor. In: The Journal of Educational Research, 58, 4, p. 163-166.

Ulbricht, K. (1986) Rükmeldung im Unterricht. Möglichkeiten und Grenzen eines Geräts. In: Medien und Technik. Series Ingenieurpädagogik, ed. by A. Melezinek, 22, Alsbach , p. 232-235

Ulbricht, K. (1983) An evaluation model in higher engineering education. In: Higher Education in Europe (UNESCO), VIII, 2. , p. 23-34.

Vehrs, Wolfgang: (1986) Nicht-verbale Erlebnisbeschreibung. Ein Verfahren zur ereignissimultanen kontinuierlichen Erfassung von Erlebnisverläufen. Göttingen-Toronto-Zürich

Winnefeld, F.(Ed.) (1967) Pädagogischer Kontakt und pädagogisches Feld. München.

Wundt, W. (1896) Grundriß der Psychologie. Leipzig, 10th edition 1911.

Wundt, W. (1900) Bemerkungen zur Theorie der Gefüle. In: Philosophische Studien, Bd. 15, S. 149-182.

LEARNER ASSESSMENT AND EVALUATION: NEW METHODS

W. Wijnen, University of Limburg, Maastricht; F. Dochy,Open University, Heerlen, The Netherlands, & I. Pychova, Silesian University, Karvina, Czechoslovakia

Assessment and evaluation has received a new dimension in recent years. First, concepts, such as centralization, competence, prior knowledge, individualization, etc. has been worked out. Different researchers and institutes have tried to make applications in relation to learner assessment.
Second, there has been a tendency towards an overall assessment prophecy. This prophecy holds that it is no longer possible to consider assessment only as a means of determining which individuals are already adapted to mainstream educational practice. A conceivable alternative is to reverse this sequence of adaptation. Rather than requiring individuals to adapt to instruction, the appropriate objective is to adapt the conditions of instruction to individuals to maximize their potential for success.
Third, these developments have been enhanced by educational technology. New possibilities, such as itembanking, computer assisted testing, interactive tests, etc. show interesting applications both in research and educational practice.

This symposium focuses at new methods for learner assessment. Different papers will stress the past weaknesses of assessment and will present several alternatives.

References

Keeves, J.P. (Ed)(1988). Educational research, methodology and measurement: an international handbook. Oxford/New York: Pergamon Press.

Dochy, F.J.R.C. (1992). Assessment of prior knowledge as a determinant for future learning: the use of knowledge state tests and knowledge profiles. Utrecht/London: Lemma B.V./Jessica Kingsley Publishers.

Dochy, F.J., Bouwens, M.R. (1992). Toetsing en toetsservicesystemen voor individuele leertrajecten in open leersituaties. Leiden: Spruyt, Van Mantgem & De Does.

LEARNER ASSESSMENT AND EVALUATION: CAN IT BE DIFFERENT?

Kari Smith, Oranim School of Education of the Kibbutz movement, Haifa University, Israel

This very short summary is only meant to be an appetizer for educators who are interested in learning more about new ideas and concepts related to evaluation and assessment within an educational setting. The purpose of the paper is therefore not to give a detailed description of the ideas presented, but rather to provide food for thought which might lead educators to examine their own ways of assessing learners.

Educators constantly assess their learners, and a mark has to be given at the end of the course whether we like it or not.

However, how do we know that the assigned mark is a fair mark which reflects what? The efforts put into the course? Class–participation? The learner's knowledge of the taught material – which does not necessarily have to be measured achievements? Personal progress? Is the mark norm–referenced or criterion referenced? The questions are many, and the answers usually vary from teacher to teacher and from learner to learner. We all have different criteria for assessing.

The process of assessing learners that will be presented in the paper is based on the following assumptions.

- The learners must be involved in the assessment process.
- The criteria for assessment must be known to the learners.
- The assessment must be based on more than one source of information (several course components).
- Each component (source of information) must be allotted a specific weight which must be known to the learners.
- The learners must be actively involved in the decision making process about what sources to draw information from and their allotted weight.

The steps on the implementation of these five assumptions will be discussed in my presentation.

ASSESSING ECONOMICAL COMPETENCE: THE OVERALL–EXAMINATION

Mien S.R. Segers, Department of Educational Development and Research, University of Limburg, The Netherlands

The education system of the Faculty of Economics and Business Administration of the University of Limburg is problem–based and has an interdisciplinary character. The Faculty offers a program which stimulates students not only to learn basic facts, methods and theories in the domain of economics. The use of knowledge and skills to solve problems within the context of every day life is considered to be essential. Reflection on and evaluation of the education system during the first 5 years of the faculty (it has been established in 1984) led this academic year to a

reorganisation of the program of the first year and a search for more valid examination methods. Beside assessing reproduction and comprehension of knowledge , quantitative skills and writing skills, a new type of examination is implemented: the Overall–Examination. The paper which is outlined in this abstract will describe different aspects as construction, correction, scoring, analyzing and related problems . It rests on the results of two pilot–studies and of the first Overall–Examination in January 1992.

The core goal of the Overall–Examination is, conform the main goals of the faculty, assessing the use of knowledge and skills in the context of interdisciplinary problems.

After two periods of eight weeks ('blocks'), during which the students dealt with a specific theme, they have two weeks self–study time. They receive a so called 'study–manual' which contains articles of scientific journals and/or relevant newspapers and some study advices. The character of the articles is different. They can be the description of a case, the report of an experiment, comments on a theory or model etc. After this period of self–study, students are assessed by the so called 'Overall–Examination'. The questions rest on the articles studied at home. They require from the student to transfer the knowledge and skills acquired during the two blocks to the articles studied. Besides questions about the material studied at home, a case handed–out during the examination is questioned in the same way as the articles studied on forehand.

The examination is constructed by a team of teachers of the different disciplines which are studied during the first year. They are responsible for the construction, correction, scoring and analysis of the results. Choices were made concerning different aspects e.g. the question format, correction methods, number of cases/articles, number of questions, structure and character of study material and advices, norm–referenced vs. criterion–referenced testing, transformation of scores to a ten point scale. These choices were based on experiences with two pilot–studies which revealed some validation problems, reliability problems and organisational constraints. They will be evaluated again after the January Overall–Examination.

References.

Segers, M.S.R., et al. (1991). The overall–test. The results of and experiences with an experiment. Maastricht: University of Limburg, Faculty of Economics and Business Administration.

Segers, M.S.R., et al. (1992). The overall–test: a second experiment involving a new assessment methodology. Maastricht: University of Limburg, Faculty of Economics andBusiness Administration.

Segers, M.S.R. (1992). The results of the analysis of the first formal overall–test. Maastricht: University of Limburg, Faculty of Economics and Business Administration.

804

INDIVIDUALIZED EXAMINATIONS AT THE DUTCH OPEN UNIVERSITY: THE SYSTEM

F. de Roode & K. Glas, Centre for Educational Technology and Innovation, Department of Examinations and Registrations, Open University, Heerlen, The Netherlands

The educational system of the Dutch Open University (Ou) is characterized by open entry (no formal educational qualifications required), open programming (modular system), and freedom of pace, place and time. the course is the central educational unit at the Ou. Students enrol for one or more courses, which can be combined into a full degree programme. A course can be completed with a formal examination and students who pass obtain study points and a certificate. In the present paper an individualized examination system is described, which complies with the open character of Ou. The system, which is described in detail in the paper, has the following setup. Item banks store the item numbers and various item data, the actual questions are stored in so–called item books. The student registers for an examination at the study centre when he or she is ready studying the course. At the examination session, the computer generated a set of item numbers for each individual student. The student looks up the questions in the item book and completes the appointed assignments. The answers to the assignments are fed into the computer and the student immediately gets a provisional result. Once the data are processed at the central cite in Heerlen, the student will get a final mark and, if the examination is passed, a certificate.
In the present paper, the various considerations which lead to developing the system will be discussed. Both practical (such as, for instance, securing the item bank and the examination situation against fraud), and technical aspects (local area network) of the implementation of the system will be described. Finally, it will be pointed out how the system is functioning in practice.

PSYCHOMETRIC ASPECTS OF ITEMBANKING: INDIVIDUALIZED EXAMINATIONS AT THE DUTCH OPEN UNIVERSITY

D. ten Brinke, F. de Roode & K. Glas, Department of Examinations and Registrations, Centre for Educational Technology and Innovation, Heerlen, The Netherlands

In the present paper, psychometric aspects of managing an individualized examination system based on an item bank, such as used at the Dutch Open University (Ou), are described. In the individualized examination system of the Ou, a student's examination consists of a randomly drawn subset of items from an item bank. Since the chance that two tests are completely identical is extremely small, it can be assumed that every student gets an unique test. As a consequence, the standard theory for the analysis of examination results, and the associated computer software, no longer applies. In this paper, the Rasch model is used for

evaluating the reliability of the items, both in difficulty of the tests and longitudinal shifts in the difficulty of the tests and the ability of the students. To perform these analyses, item and person parameters of the Rasch model are estimated using a conditional maximum likelihood (CML) procedure. Further, model fit is evaluated using likelihood ratio tests.
Several examples using data of Ou examinations will be presented. The item banks consist of 208 to 837 items, test length varies from 20 to 80 items. The examples include item analyses and estimates of the percentage of miss–classifications due to the test administration procedure.
Finally, some suggestions with respect to future developments of the system will be made, such as minimizing the fluctuation of the difficulty of the tests using optimal item selection.

ASSESSMENT STRATEGIES FOR DIRECTING FURTHER LEARNING IN MATHEMATICS

J. Izard, Australian Council for Educational Research, Australia, & C. Haines, City University, London, United Kingdom

Those who assess have to cope with a number of conflicting demands. On one hand there is the desire to reduce the time devoted to the assessment task to the absolute minimum and to make the task as simple as possible. Conversely there is the necessity to ensure that the assessment is as comprehensive as possible to meet accountability requirements – fairness to each student so that they are rewarded in accord with their genuine knowledge and skill, providing students with information to support their own learning, trustworthy certification to reassure the community and informative for those who will use the assessments as part of the selection process for entry to later stages of education.
Many examinations are restricted to pencil–and–paper tasks, distorting the assessment by failing to sample tasks where there are multiple (correct) solutions, where there are a diversity of problem–solving approaches, or where the skills cannot be demonstrated easily in pencil–and–paper format (de Lange & Verhage, 1990; Izard, 1990). Since students, teachers and administrators consider examination results are important, they concentrate on teaching to the anticipated examination questions rather than to the achievement itself (Madaus, 1988). Assessment strategies are needed to cater for a wider range of mathematical tasks.
This paper describes a study in which comprehensive descriptors of achievement were developed to assess complex behaviour (as observed in mathematics projects undertaken at British universities and polytechnics) to recognise high achievement and to reward curriculum–intended achievement. Analyses of non–dichotomous data from trials of the scheme are presented and interpreted in terms of the consistency and precision of assessing candidate achievement, the stringency of the examiners, and the differing demands of particular course requirements.

PROBLEMS IN MEASURING THE RELIABILITY OF NATIONAL CURRICULUM ASSESSMENT IN ENGLAND AND WALES

Ian Schagen, National Foundation for Educational Research in England and Wales, United Kingdom

Summary

The 1988 Education Reform Act set up the structure of the National Curriculum and its assessment for England and Wales. The points in a child's career at which National Curriculum Assessment (NCA) is to be carried out are specified as the end of "key stages" in their progress through the maintained school system. There are 4 such key stages, and the ends of them correspond respectively to the approximate ages of 7, 11, 14 and 16. Individual assessments are quoted in terms of "level attained", an integer having values from 1 to 10 (plus "W", working towards level 1, roughly equivalent to 0). For each Attainment Target (AT) the level attained is determined by a set of "Statements of Attainment" (SoAs), one or more for each level of the AT.

Overall, NCA involves two strands or modes of assessment. The first is "Teacher Assessment" (TA), the on-going process by which class teachers record the progress of their pupils against the ATs of the core subjects, and at the end of each Key Stage formally note down the levels achieved on each. The second mode of assessment makes use of a "Standard Assessment Task", a predefined set of activities which monitors performance on a subset of the ATs in the National Curriculum. The task produces information on individual SoAs, which are then aggregated to give a level for the AT.

In this paper we shall concentrate on results obtained from the Key Stage 1 pilot tests (for pupils aged approximately 7) carried out in May 1990. Each task comprised a set of activities, with each activity containing a mixture of items addressing SoAs from a variety of ATs. The activities were thus "cross-curricular", and it was not possible to derive the AT levels directly from the data as it was collected. The data from all pupils involved in the pilot was collected and entered to a computer which carried out the "aggregation" procedure.

One of the questions to be answered about NCA data is the extent to which assessments on a given Attainment Target (AT) may be considered to be measurements of a single ability or continuum. The program TESTFACT is designed to carry out categorical factor analysis on item response data, based on a multidimensional IRT model, and this program was applied to selected datasets from the 1990 pilot. TESTFACT found problems in dealing with missing values,

which tend to occur in this data whenever teachers abandonned items or whole activities due to lack of time or for other reasons. A solution which was adopted was to "impute" values for the missing data, using the aggregated result for the non-missing items on each AT. Two separate runs were therefore made with each data set for each AT, with the missing values either treated as incorrect, or imputed.

For each TESTFACT run we attempted to extract 3 principal factors and apply VARIMAX rotation to group items into components. For the raw data the factors extracted corresponded to the various activities making up the SAT. When imputed values were substituted for the missing data, the activity- based structure tended to disappear. The factors extracted seemed to relate more to levels than to anything else. This could mean that items at the same level are more "clumped" than would be expected on TESTFACT's underlying IRT model, but this could be an artefact of the way in which missing values were imputed. From a study of the latent roots from each factor extraction it was clear that the first factor is dominant, and further factors are very much smaller, especially with the imputed data. Even with raw data this is the case, implying that the underlying structure of each AT is essentially one-dimensional.

Given a set of outcomes for the items addressing a particular AT, measures have been developed which help to answer questions about the consistency of items and responses. The first measure, of "orderedness" (G), tests the extent to which the item facilities match the assumed hierarchy of AT levels. In other words, do items at lower levels have higher facilities than those at higher levels, or do certain items appear to be operating at the wrong level? The second measure, of "consistency" (C), tests the extent to which individual pupils have scale-type responses. In other words, do they operate consistently at a distinct level, passing items below that level and failing items above it? Both measures operate on a scale from -1 to +1, with zero indicating random behaviour. Values calculated for the ATs assessed in the 1990 pilot tests showed good agreement with more detailed analysis of the operation of the items.

The coefficients G and C which have been developed are useful measures which describe the behaviour of the data in terms of consistency and orderedness; however, we need to have a better measure which describes the extent to which the assessment as a whole gives an accurate picture of the pupil's underlying position relative to the AT. To develop such measures we need to have a conceptual model of the pupil's responses to the items in the assessment task. We shall describe a "base model" which will attempt to encapsulate those aspects of the NCA data which are distinctive, without becoming over-complex. This base model makes a number of simplifying assumptions in order to achieve a theoretical framework which can be manipulated in a suitable way. These assumptions are:

- For the AT of interest, each pupil has an underlying "attainment level" which can be considered to be a real-valued random variable, Normally distributed with mean and standard deviation depending on the individual pupil. Realisations of this random variable can be regarded as real numbers lying on the same scale as the NC levels.
- Each item which addresses the given AT has a nominal level, which takes on an integer value (from 1 to 3 for KS1), and a true difficulty" parameter, which is real-

valued. If an item has its difficulty equal to its nominal level, we shall say it is "on level".

- The probability that a pupil passes a particular item is given by the Standard Normal integral, combining pupil and item parameters. Given the pupil and item parameters, the probabilities of passing different items are independent.

Using this base model, we may estimate its model parameters from actual NCA data, and hence obtain information on the underlying variability of the items and the pupils. Given a set of "pupil profiles" (combinations of item responses obtained by one or more pupils), it is relatively simple to write down an expression for the total log-likelihood in terms of the model parameters. In principle we then just need to find values of the pupil parameters and the item difficulties to maximise this. In practice the number of parameters involved is excessive for a direct multidimensional optimisation to be feasible, so a two-stage approach is used, successively estimating the pupil parameters with the item difficulties fixed, and vice versa, until convergence is achieved.As well as the item difficulties, our base model provides us with 3 parameters relating to the pupils: the mean and standard deviation of the pupil abilities, and the within-pupil standard deviation. The latter can be regarded as a measure of the reliability of the assessment, in units equivalent to levels. Low values imply that pupils are consistent in their responses to items, and high values indicate very inconsistent responses to the items, and thus an unreliable assessment.

The within-pupil standard deviation by itself is not, however, a sufficient indicator of reliability. It is not intuitively clear what its values represent - is a value of 0.3 good, or poor? It is possible to get incorrect assessments with low values of Â if the item difficulties are badly distributed, i.e. if there are several "out-of-level" items. Furthermore, the assessment may be more reliable for some levels than others - it may distinguish well between level 1 pupils and the rest, but be poor at identifying level 3 pupils.For all these reasons our measure of reliability or its equivalent needs to be more sophisticated than can be provided by a unique number. The suggestion is to compute the probability of a correct assessment at a given level, dependant on the item difficulties, the within-pupil standard deviation and the pupil ability distribution. We also need to take into account the aggregation rule in force for the 1990 NCA (i.e. if there are 1 or 2 items at a level, all must be passed; for 3 or more items, one may be failed and the level still awarded).

Three probabilities may be defined for any level l : of correctly passing (R1[l]), correctly failing (R2[l]), and of an overall correct result (R3[l]). These define a set of measures for the assessment of an AT based on the fitted model. Software to carry out this analysis has been developed and run for various datasets from the 1990 pilot.The combined probability at level 2 might be regarded as the clearest reliability measure for a given AT and dataset. For Ma 2 (Number and number notation) it was found to be around 0.86, and it varied around about 0.80 for En 2 (Reading). Sc 1 (Exploration of Science) was more variable, with values from 0.47 to 0.77. The dataset with the largest number of items for Sc 1 (TG) also had the lowest reliability and the most "off-level" items.

The measures derived here combine information about the pupils' responses to the items, the structure of the AT and the aggregation method used to give a measure of the overall "reliability" of the NCA, in terms of the probability of pupils being correctly classified at different levels. In this way they can contribute to the

803

design and evaluation of more reliable National Curriculum Assessment procedures.
(Funding for this project was provided by the Economic and Social Research Council under Research Grant no. R000233076).

References

DES and WELSH OFFICE (1987) National Curriculum: Task Group on Assessment and Testing - A Report, HMSO, London.

Whetton C. et al (1991) The Pilot Study of Standard Assessment Tasks for Key Stage 1, SEAC, London.

NATIONAL CURRICULUM ASSESSMENT IN ENGLAND AND WALES

Marian Sainsbury, National Foundation for Educational Research, United Kingdom

The first national assessment of seven-year-old children in England and Wales took place in the Summer of 1991. The assessment had two elements: continuous assessment by teachers; and standard assessment tasks. These tasks were devised by a team from the National Foundation for Educational Research, working with Bishop Grosseteste College. The team also carried out an evaluation of the national assessment system, which consisted of a questionnaire survey, case studies and an analysis of pupils' results.

In 1988, the British government introduced, for the first time, a National Curriculum for all pupils between the ages of five and 16 years in England and Wales. The National Curriculum was structured by subject, and included the `core subjects' of English, mathematics and science, and `foundation subjects' of technology, history, geography, art, music, physical education and a modern foreign language. (The modern foreign language applied only to pupils aged 11 and over.) Each subject was subdivided into a number of attainment targets. The National Curriculum was introduced gradually, starting with the core subjects.

A Task Group on Assessment and Testing was set up to advise the government on the assessment arrangements for these subjects. They advised a criterion-referenced system of ten levels of attainment within each subject. The levels should cover the entire range of attainment from five to 16 years of age. Children would progress through the levels at different speeds, but each level would represent about two years' progress for the average child.

The Task Group on Assessment and Testing advised that pupils should be subjected to formal assessment at the ages of seven, 11, 14 and 16. This should consist of a summary of the teacher's continuous assessments and the results of standard assessment tasks. These tasks were envisaged as an entirely new form of assessment, with a variety of modes of presentation, operation and response. The assessment system was designed to fulfil five purposes: it was to be formative, summative, evaluative and diagnostic, and to bring about professional development for teachers.

The standard assessment tasks were administered for the first time in 1991, to about 600,000 pupils throughout England and Wales. Welsh-speaking children took standard tasks in the Welsh language. The tasks covered nine attainment targets: reading, writing, spelling, handwriting, number, using and applying mathematics, exploration of science, and one further topic in both mathematics and science, from a constrained choice. The tasks were performance assessments. Teachers were given instructions to set up activities that reflected normal classroom practice. They were given criteria for observing pupils' oral, practical and written responses, and a format for recording their assessments.
The nature of the standard tasks varied according to the subject matter being assessed. Two examples may illustrate this variation. To assess number, pupils were given a worksheet showing some 'shopping' calculations for them to complete. By contrast, the assessment of exploration of science consisted of a practical activity. Children were given a range of objects, and were asked to weigh each one, and then make a prediction about whether it would float or sink if placed in water. They went on to test out these predictions, while their teacher observed such things as their formulation of hypotheses, their ability to record their results in a chart, and to draw conclusions from their investigations.
In order to administer these tasks, teachers had to organise their classes so that they could spend an extended period of time with a small group of children, observing and recording their responses. A training programme was instigated for teachers, to address matters of organisation, observation and recording. The training was carried out by advisory staff from local education authorities, who also made visits to schools and advised upon the interpretation of the criteria. Two main issues emerged from the evaluation of the standard tasks: classroom organisation, and consistency of judgements.
In general, the most successful pattern of organisation was displayed by those teachers who spread the tasks over the maximum permitted period, six weeks, and integrated them into the normal curriculum. All teachers, however, reported a greatly increased workload as a result of the standard assessment tasks. Complaints focused, in particular, on the practical activities such as the science task described above. The teachers' workload received wide coverage in the national press and was the focus of attention from government ministers.
The administration of the standard assessment tasks also represented the first national application of the criteria defining each level of National Curriculum subjects. These criteria had been devised by groups of subject experts and varied in their precision and explicitness. Despite training, teachers were observed to vary in their interpretation of some of the less precise criteria, especially when assessing practical or oral responses. The standard assessment tasks were an important element in establishing consistency of judgement amongst teachers, but much work remained to be done by further training.
Although early reactions to the tasks tended to stress the negative responses to the extra workload, later observations and interviews revealed a number of beneficial effects of the standard assessment tasks. Headteachers and trainers remarked upon the professional development that resulted, in terms of teachers' improved organisational and observational skills. Teachers' familiarity with the requirements of the National Curriculum was enhanced and their capabilities in carrying out classroom assessments were improved. (See Gipps, C. et al,

National Assessment at Seven: some emerging themes Bedford Way, Kogan Page, London, 1992.)

As a result of the evaluation of the 1991 assessment process, a number of changes were made to the specification for the 1992 standard tasks. Only seven attainment targets were addressed, a reduction of two, and the time-consuming practical activities, exploration of science and using and applying mathematics, were excluded. A greater emphasis was placed upon worksheets that could be completed by larger numbers of pupils, working simultaneously. Nevertheless, teachers retained the option of administering the tasks in an informal way and integrating them into normal classroom routines. A longer administration period, of twelve weeks, was permitted. Preliminary indications from the 1992 evaluation indicate that teachers were less apprehensive about their workload than in 1991, and that the reduction in subject-matter and the lengthening of the administration period were welcome.

The development and evaluation of the 1991 and 1992 standard assessment tasks were carried out under contract to the School Examinations and Assessment Council.

812

THE DUTCH NATIONAL ASSESSMENT PROGRAM IN EDUCATION: AIMS AND SCOPE

J.M. Wijnstra, National Institute for Educational Measurement (Cito), Arnhem, The Netherlands

Introduction

Last year the data collection was finished for the first five year cycle of the Dutch National Assessment Program in Education. In this cycle the whole curriculum of elementary education was covered with assessments at the end of elementary education, and additional assessments in the fields of mathematics and Dutch language about midway through the elementary school.

The general aim of the program is to provide concrete data for a well–founded discussion of the contents and level of education. Each assessment starts with the development of a domain description. This structured list of objectives and topics in a subject, which is not restricted to a (yet non–existing) core curriculum, serves as the basis for the instrument construction. The assessments are conducted on samples of the elementary school population. By means of questionnaires, data are collected on the organization and the curriculum in use, i.e., the topics actually taught. The students take tests and exercises to determine what they actually know and are able to do. In this way a picture emerges of the condition of education, with differentiations regarding background characteristics of schools and students.

Project staff is not allowed to draw any conclusions about the quality of education. What constitutes an acceptable level of education should be left to the relevant authorities. The project is only providing a mirror. To get an idea of what interested and concerned parties expect and desire as outcomes of education, teachers and parents are consulted. The results of these consultations can serve as a point of reference and are therefore presented in the reports, together with the comments of some experts in the field.

In the contributions to the symposium several aspects of the program will be highlighted with illustrations from various assessments:

- The operational curriculum in science and social studies education;
- Problem solving methods in mathematics;
- A survey of students' attitudes on learning English;
- How to interpret the assessment outcomes: expectations and desires.

THE OPERATIONAL CURRICULUM IN SCIENCE AND SOCIAL STUDIES EDUCATION

J. van Diggele, & J van Weerden, National Institute for Educational Measurement (Cito), Arnhem, The Netherlands

In this contribution the focus is on the route from domain description to the analysis of data concerning the operational (actually taught) curriculum. For this goal science and social studies education will be used as an example. The data collection took place in 1990 and 1991.
The domain description for science and social studies is rather extensive and specific. The domain is divided into six parts, which in turn are subdivided into topics and aspects. It includes topics that are common practice and belong to a more traditional approach as well as topics that have a more innovative character like pollution and health education.
As part of the data collection the teachers of the three final grades filled out a questionnaire about the way they organize the curriculum and teach the subject, including an inventory of topics taught. The collected data about the actually taught contents are used for the description of the state of the art. Some of the variables are also used to explore relations between student performance and the operational curriculum.
The results show that the average school spends about five hours a week on science and social studies on a regular basis. In addition, some time is spent on special projects. Compared to other subjects also the variation in curriculum materials is quite large. These differences reflect the fact that science and social studies are changing subjects. The data about the contents actually taught show that 75% of the topics in the domain description is taught in at least 75% of the schools. Lower percentages are found especially in the field of physics and with respect to social–cultural and economic concepts.

PROBLEM SOLVING METHODS IN MATHEMATICS

J. Bokhove, J-M. Kraemer, & J. Janssen, National Institute for Educational Measurement (Cito), Arnhem, The Netherlands

In the 1987 mathematics assessment an inventory was made of the way students solve problems. The problems were 30 items taken from 8 of the 27 scales used in the assessment. They were presented to a random sample of students, asking them to explain how they reached their answer.
The reasons for studying the way in which a solution is reached, were twofold. In the first place, paper and pencil tests cannot always adequately show whether a certain goal is reached. In the second place, the individual's ways to reach a solution are heavily accented in modern approaches to realistic mathematics. However, most of the available observations are incidental and there is great need for more systematic research on this subject, especially on a larger scale.
In this contribution we report on two topics:

- the conclusions that can be drawn from an individual evaluation of mental arithmetic;
- the kind of errors made when solving simple context problems.

In mental arithmetic and estimation many pupils follow unwanted lines of reasoning to get at a solution. Instead of using easy and accessible ways to solve a problem, they use inadequate mental calculations. Instead of starting by rounding off the numbers, they first calculate exactly and they round off the result. In general three types of errors can be distinguished in solving context problems: erors of interpretation, conceptualization, and performance. Interpretation errors are based on a mistaken translation of the problem into a mathematical problem. Conceptualization errors are based on errors in the mathematical concepts. Performance errors are purely arithmetical errors as a result of insufficient control of basic operations or mistakes in the procedures.

HOW TO INTERPRET THE ASSESSMENT OUTCOMES: EXPECTATIONS AND DESIRES

J. Sijtstra, National Institute for Educational Measurement (Cito), Arnhem, The Netherlands

The general aim of of the assessment program is to describe the actual level of performance in various school subjects related to the contents teachers offer their students. One of the problems is how to interpret the data. What can be said of the performance levels at the end of elementary education? Project staff should leave conclusions about the quality of education to the relevant authorities.
Nevertheless, as a way of initiating the discussion about the level and contents of education, some tools for the interpretation of the results are available. Amongst these are the results of the consultation of parties who are directly interested in and concerned with a certain school subject. Without prior knowlegde of the assessement outcomes, these parties are invited to express which performance levels they expect and desire.
In this contribution this kind of research will be exposed, illustrated with some cases drawn from the assessment of Dutch language proficiency. The main questions the respondents have to answer are:

- Which items do you expect the average student to master well (with 80% chance of succes) at the end of elementary education?
- Which items do you desire the average student to master (to the same degree)?

In the case of integrated speaking and writing assignments the question is:

- What percentage of students do you expect to produce a text of a well defined quality, respectively what percentage do you desire?

The design and the results of this consultation are presented. The estimations of the average student's level of listening and reading comprehension are quite adequate. The actual levels of speaking and writing are overestimated, however, and the desired level is even higer than the expected level. Furthermore, the several parties consulted differ in their expectations and desires.

A SURVEY OF STUDENTS' ATTITUDES ON LEARNING ENGLISH

M. P. Vinjé, National Institute for Educational Measurement (Cito), Arnhem, The Netherlands

The 1991 assessment of English included student achievement in speaking, listening and reading comprehension, as well as students' attitudes toward learning this language. In order to measure these attitudes a questionnaire was designed consisting of 28 questions on topics such as: difficulty, attraction, importance, readiness to speak English, and actual use. The questionnaire was given to 553 pupils, randomly sampled from 188 schools in the Netherlands. Factor and cluster analysis of the answers show several factors. The most important factor can be labeled as 'self evaluation of one's own capacities in English and willingness to use them'. Other factors are related to the comparison between the school subject English and other subjects, such as geography, history, Dutch and mathematics. The results indicate that students perceive themselves as quite capable of speaking and understanding English, and willing to do so. In comparison with other subjects, English is regarded favourably.
In the contribution these results will be discussed in relation to some aspects of the students' actually observed audacity and oral performance in English, as well as some of their background characteristics.

816

TOWARDS A STUDENT MONITORING SYSTEM IN PRIMARY SCHOOL

J van Bergen, L. van den Bosch, T. Eggen, R. Engelen, P. Gillijns, F. Kamphuis, & F. Moelands, National Institute for Educational Measurement (CITO), Arnhem, The Netherlands

Many recent pedagogical developments in Dutch education share the element of the necessity of obtaining an adequate impression of the progress made by each individual student and by groups of students. The Student Monitoring System that the Institute for Educational Measurement in The Netherlands is developing will be presented here as a teaching aid for monitoring systematically the progress made by individual students and by groups of students over a number of years. It creates the opportunity to determine twice a year whether a student is making enough progress in the basic skills of oral and written language, reading, arithmetic, social sciences and information processing.

By applying item–response–technology, scores on successive tests can be transformed to one and the same underlying scale. In fact, this enables the teacher and the school to monitor a student's progress. In order to be able to decide whether a student's progress is satisfactory or not, the results can be compared with those of different groups of students from across the country.

Scoring and further processing of the tests, as well as registration of the results, can be done manually. In addition, a software package will be developed for computerizing the scoring and processing of the tests, and the registration of the results.

In this symposium the outline for the Student Monitoring System will be presented and the applied measurement technology will be highlighted. How the general characteristics of the Student Monitoring System have been worked out in detail for a test, will be presented by means of the example of the Scale Progress in Spelling Ability. Finally, the development procedures, operation and possibilities of a computer program belonging to the system will be presented.

In the first presentation the outline for the Student Monitoring System, designed at the Dutch Institute for Educational Measurement, is presented. The system is an aid for the teacher and the school for monitoring the progress made by students and the quality of the education offered, thus enabling them to make well–founded decisions.

In its final form the system consists of a coherent set of measuring instruments for long term evaluation of educational progress. The instruments cover the basic skills: oral and written language, reading, arithmetic, social studies and information processing. Twice a year, the progress of the student is measured on one or more

parts of the curriculum, with the intention to check whether a students's individual growth is satisfactory, in comparison with previous occasions and in comparison with relevant (sub)–populations.
An attempt has been made to avoid the teacher from being empty–handed, when he has noticed that a pupil has made insufficient progress or has an insufficient ability in one or another subject matter. Whenever possible the manuals will contain didactic instructions and exercises that can be used in remediation, references to teaching materials, textbooks or remediating material, or even additional tests for measuring relevant sub–skills.
After a discussion on the feasibility of systems monitoring the educational outcomes of children, the basic characteristics of the Student Monitoring System will be introduced. Next we will present the premises and the content of the Student Monitoring System, and describe the longitudinal research design and the applied measurement technology in global terms. Subsequently we will discuss the scoring and further processing of the tests, the registration and interpretation of the results and the educational procedure in which the system should be used. Finally, the use of the proposed assessment system in educational practice will be discussed.
As an example of the progress scales in the student monitoring system, the orthography or spelling test will be presented in the second presentation. The general aims in this test are:

- to determine whether a child masters the spelling subjects to a certain level at a certain moment;
- to compare the spelling proficiency of children with those of relevant national reference groups;
- to determine whether a pupil exhibits sufficient progress compared to a previous assessment.

Certain aspects of the test will be explained in order to show how the aims in the test are accomplished. Attention will be payed to the content of the test, tailor–testing and the administration and interpretation of test results. One of the complaints of users of tests often concerns the fact that no indications are given about educational measures that can be taken to help children with poor test results. We have tried to meet the wishes of the users by describing a way of locating gaps in the spelling proficiency. When teachers know with what kind of words children have specific writing problems, they can make use of a special quire that is added to the Scale Progress in Spelling Ability and which contains exercises and references to methods.
In the third presentation the measurement technologie will be highlighted. The task of characterizing growth in level of attainment within an individual student is far from easy. One reason for this is the typical unreliability of educational measures. So, more structural information concerning the student, like measurements on previous occasions of the same and other skills and background information, could be used to reduce this uncertainty.
In our approach, the measurement problem and the structural problem are separated. Measurement models, mostly Item–Response–models, are applied to get a good estimate of a student's ability. In this stage, we have to deal with issues like the (vertical) equating of tests in a longitudinal incomplete setting and the statistical testing of the applied models. In the structural part, models are used

818

which utilize all the available information, i.e. a growth model is specified (e.g. State Space models). Combining the information from the structural and the measurement model enables us to update the estimates of student's ability from the measurement models by filtered or smoothed estimates (Kalman filter). In this presentation the methodology will be elaborated and some preliminary results will be presented.

In the final presentation the development procedures, operation and possibilities of a computer program belonging to the system will be presented. For the development of the computerized part of the Student Monitoring System the Institute for Educational Measurement cooperates with the staff of the Project for Implementation of New Technologies in the Netherlands (PRINT).

The software is built for specified hardware to be used in all Dutch primary schools. Both the hard– and software–specifications are prescribed by the Dutch Ministry of Education. The software–specifications include a so called WIMP's–environment. Specific procedures for software–development are prescribed by PRINT. In the case of the software for the Student Monitoring System the following phases can be distinguished: project–proposal, contracting, project–plan including feasibility–study, educational design, specification, technical implementation, acceptance and fieldtesting and finally revision and delivery.

The software–program for the Student Monitoring System consists of three main parts:

- data–input (which contains possibilities for input of testresults);
- data–processing (scoring student answers, weighting, analyzing , transforming and combining of scores);
- data–output (reporting on the screen and on paper; interactively and in batchmode).

In this presentation these three subjects will be elaborated and during the presentation a demo of the program will be shown.

CHAPTER 6

TEACHING AND TEACHER TRAINING

820

TEACHING AND TEACHER TRAINING

Jules Pieters, Department of Education, University of Twente, Enschede; Bernadette van Hout-Wolters, Graduate School of Teaching and Learning, University of Amsterdam, Amsterdam, The Netherlands

Introduction

Teaching is the central process of education. It pertains to several activities and processes intended to lead to learning. Brophy & Alleman (1991) distinguish the following roles and functions in teaching: management in the classroom and motivating students, presenting information and demonstrating procedures, asking questions and engaging students in content-related discourse, developing understandings through learning activities, providing for practice and application opportunities through additional activities and assignments, structuring activities and assignments for students and scaffolding their progress on them, and evaluating learning.

Research directed to teaching is involved in teachers' behaviour and its relation to student outcomes. Recent research on teaching extends attention to cognitions of teachers. Models have been built that reflect teachers' cognitions related to the planning of the teaching process. But building all encompassing models of teaching has been severely criticized. To Shulman (1986) it seems unlikely that any theoretical frame can encompass the diversity of sites, events, facts, and principles that cross all those levels. In spite of these criticisms research brought about several empirically founded indications for models of teaching.

In all these models, activities and processes teaching are arranged around the kernel actors: teacher and students. Many variables are common to these models: (1) perceptual and cognitive processes of teachers, triggering (2) teacher actions, facilitating (3) students' perceptual and cognitive processes, directing towards (4) students' actions, that bring about (5) students' learning outcomes.

Four sub–themes related to research on teaching, the conditions of teaching, the teaching profession, and the preparation and development of teachers have been defined on the basis of the papers submitted to the ECER conference:

- Nature of teaching and teaching to learn;
- Beliefs, attitudes, and cognitions of teachers;
- Teacher education and staff development;
- Teacher training in international perspective.

The nature of teaching and teaching to learn includes new insights in teacher knowledge structures related to effective teaching. The dynamic nature of these knowledge structures represent the development of professional competence.
Teachers' beliefs, attitudes, and cognitions will be discussed within a recent framework of situated cognition and cognitive apprenticeship.
Teacher education and staff development will be approached from several positions taken by subject–matter teaching, various professional fields, training technologies, and collaborative training design.
One of the main advantages of bringing together researchers from many European countries is facilitating discussions about national teacher education policies and teacher training methods. Teacher training in international perspective will be such a meeting place.
Within the four sub–themes various research methods and measurement techniques will be presented and discussed. New qualitative techniques are advocated by researchers from various fields and areas. Due to scarce empirical evidence, these new techniques lack reliability and generalizability.
Another encompassing feature of recent research on teacher thinking and teacher training is constructivism. Although its wording itself is not apparent in many papers presented on this conference, its influence is unmistakably present.

References

Brophy, J., & Alleman, J. (1991). Activities as instructional tools: A framework for analysis and evaluation. Educational Researcher, 20(4), 9–23.

Shulman, L. S. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. In M. C. Wittrock (Ed.). Handbook of research on teaching. 3rd Edition. New York: MacMillan.

The papers in this theme are clustered into the following categories:

- Nature of teaching and teaching to learn (AA);
- Beliefs, attitudes, and cognitions of teachers (BB);
- Teacher education and staff development (CC);
- Teacher training in international perspective (DD).

You will find the abbreviations in the first column of the program.

THE PROFESSIONAL DEVELOPMENT OF TEACHERS IN HEALTH EDUCATION: ISSUES IN THE DELIVERY OF INSERVICE PROVISION IN THE UNITED KINGDOM

Gaye Heathcote, Health Research and Development Unit, Crewe and Alsager College of Higher Education, United Kingdom

The paper describes and analyses a research, development and evaluation project designed to identify and respond to teachers' in–service professional requirements in Health Education. The project was funded by the Health Education Authority in recognition of the absence of any nationally agreed policy or major initiatives in this area and of the growing pressures on teachers to (i) lead health education initiatives in their schools (ii) forge closer health–promoting links with the community (iii) be conversant with the nature and work of other health professionals in the field (iv) implement the National Curriculum through subject–based and "whole–school" approaches.

Following the establishment of a Working Party in March 1986 to consider this problem and to explore the possibility of in–service course provision for teachers, questionnaires were sent to all Local Education Authority Advisers with a responsibility for Health Education and for allied areas of the curriculum (Pastoral Care, Equal Opportunities, Personal Education/Development, Life Skills etc) to elicit views on the extent of need and on the desired aims, nature and organisation of course–based in–service provision. Questionnaire responses indicated a clear need for the development of a set of national guide-lines and a curriculum framework within which courses could be planned in a way that responded to specific local needs of teachers in all types of educational establishments (primary, secondary and post–school education).

Further research was then designed and carried out in two phases between 1987 and 1990. The first phase involved extensive survey studies of existing practice, current provision and future requirements for teachers of Health Education using questionnaires and in–depth interview techniques. The research design had to tackle a number of problems concerning the definition of `Health Education', the scope of the term `in–service provision', the complexity of existing in–service provision at local level (which involved a large number of health and education related `agents' and `agencies') and access to teachers' and health educators' views. There was also the problem of sampling because the human and financial resource constraints imposed on the project precluded an extensive national survey. A decision was therefore taken to conduct the survey studies in selected

areas and to identify a set of criteria for selection purposes which would ensure a 'mix' of characteristics within the sample. These were (i) rural versus urban (ii) high ethnic population versus high white population (iii) new town versus inner city (i) affluent versus deprived. An additional consideration was to ensure that no large geographical area of England was left uncovered.

The survey provided much detailed information about the course participants, their professional needs, their experience of in–service provision in Health Education and their views on its effectiveness. It also provided a 'mapping' of the nature, content, organisation and outcomes of existing provision, and perceptions of course providers. The main findings were:

- The initial training of many teachers had been inadequate to respond to the current demands of school–based Health Education. However, attendance on in–service courses posed difficulties for teachers because of managers' reluctance to release them for training. The comparatively low status of Health Education in relation to other curriculum areas exacerbated these problems.
- Large numbers of different 'agencies' at local level offered a diversity of short in–service courses on topical issues such as smoking, alcohol, drug use among young people, stress reduction, child abuse, safety, AIDS and sexuality, pregnancy. These varied considerably in their length, content, organisation, recruitment, attendance modes, quality and availability. All could claim to react flexibly to teachers' immediate perceived needs, but none was able to provide a coherent, structured and progressive pattern of professional development. A small number of longer in–service certificated courses was offered in some institutions of higher education but these tended to be oriented towards health professionals other than teachers. Moreover, teachers often experienced structural, financial and logistic problems in relation to course attendance. Both these sets of factors contributed to the low recruitment and marginalization of teachers.
- Finally, the concept of 'Health Education' was often misunderstood by policy–makers and budget–holders. School–based health education in Britain seeks to promote physical, mental, emotional and spiritual health by using a variety of pupil–centred teaching methodologies. Delivered from many different subject bases, eg., Physical Education, Science, Humanities, Religious Education, as well as through the pastoral care system, it offers not only 'facts' about healthy lifestyles but creates opportunities for the development of autonomy and self–empowerment, for the practice of decision making skills and for the sensitive exploration of values, meanings, perceptions, opinions and feelings that are associated with health practices. These principles were not always understood by those in power.

This first phase of the project enabled recommendations to be made about the nature and content of a new national programme that was capable of confronting such problems. The research findings indicated that such a programme should eg..

- Recognise and accord credit for teachers' prior in–service training and experience whilst seeking to extend their experience beyond school–based health education to embrace a wider vision of the nature, content and possibilities of health promotion. This would focus on an exchange of views and experiences between teachers and other types of professionals involved in health promotion (eg doctors,

dentists, environmental health officers, pharmacists, social workers, dieticians, chiropodists, physiotherapists, nurses, health visitors).

- Equip teachers to undertake in–service work with colleagues, school governors and the local community, and organise collaborative activity with colleagues from other schools and other communities.
- Ensure that teachers have the opportunity to implement and extend the skills they acquire while on in–service programmes by working closely with resource managers at all levels.
- Involve teachers in the teaching methodologies, content and evaluationstrategies which they conventionally use with their pupils in school–based Health Education. This means ensuring that the in–service programme is truly `person–centred', that is embraces principles of empowerment and facilitation, recognises and shares the experience and expertise of the group, and emphasises doing, experiencing and skill–building in a context which values the dignity of individuals and the growth of self confidence and control over one's life and health `career'.

The second phase of the research developed a curriculum and organisational framework predicated on these findings and funded two courses in different parts of England to test the feasibility of the proposals. The course which the author developed and piloted in her own institution consisted of four phases:

- A pre–course experience in which teachers undertook community based health promotion. This involved `shadowing', observing or working with other health professionals `in the field'.
- Open learning (involving text, tape and video), supported by telephone tutorials and self–help work (`paired' or small group with other teachers attending the course from the same locality), workshop attendance and residential weekends (where practical group work and communication skills as well as confidence–building, relaxation, curriculum planning, implementation and evaluation strategies were developed).
- Two school–based phases during which the teacher (i) instigated awareness–raising and innovation in Health Education through work with pupils and (ii) involved colleagues in a `whole school approach'. This meant working collaboratively to develop school–based health education policies and to plan, organise and co–ordinate the contributions of others to the goals of health promotion.

This pilot was evaluated `in–house' using an independent researcher, questionnaire responses, in–depth interviews with course participants, external examiners' reports and self–assessment techniques. It was also evaluated by an external examiner who collected feedback from the two institutions involved in the pilot. The evaluation indicated increased confidence on the part of teachers to lead and co–ordinate in–service Health Education, increased awareness and knowledge of health education issues and enhanced professional skills.

Reference

Heathcote, G. (1989). In–Service Provision for Teachers of Health Education. London: Health Education Authority

DEVELOPMENT OF AN INSTRUMENT FOR ASSESSING CLASSROOM PERFORMANCE IN A FACULTY OF MEDICINE

Th.J. ten Cate, Educational Development, Faculty of Medicine;
J.A.L. van den Bijtel, Centre for Educational Research SCO,
University of Amsterdam, Amsterdam, The Netherlands

Frame of reference

Assessment and training of teachers in universities is only recently a major topic of interest in the Netherlands. Historically teachers of scientific education are selected upon other qualities than educational skills. Research merits prevail and, in medical education, quality of patient care. Stimulated by government policy (2) assessment of quality of higher education has become current (3). Nevertheless structural assessment of didactical skills of individual teachers in higher education is still unusual for a number of reasons. Instruments used have often a weak theoretical basis and are incidentally applied with little structural consequences. Formal aspects of these consequences have been subject of discussion (7, 8).
In 1989 the Board of the Faculty of Medicine of the University of Amsterdam adopted the initiative to develop a faculty wide system to regularly assess and train teachers, as a part of the faculty policy. A three year project was designed to prepare assessment instruments, to develop a teacher training program and to develop adequate personnel policies pertaining to assessment and training. This paper reports on the first project year, in which a pilot study was carried out to develop a set of assessment instruments for teachers in the preclinical period of medical education (5).

Theory

Process–oriented instruction theory formed the basis for the development of the instrument. Vermunt has built a framework in which selfregulation–of–learning theory and instructional theory are combined into this process–oriented instruction theory (6). Effective instruction should, in his view, match learning activities as closely as possible. The following process–oriented approach to instruction served as a basis for the construction of the assessment instrument.
While studying students develop cognitive, affective and metacognitive activities. In other words: they process factual content–information, they maintain motivation to study and they control their learning activities and strategies. These three groups of self–regulatory activities can be divided in different specific sub–categories. Vermunt identifies eight categories of cognitive processing activities,

eight categories of affective (i.e. motivational) activities and another eight metacognitive regulation activities, based on sustantial research.
Teaching in higher education has traditionally been limited to presentation of (oral and written) information and has only gradually shifted towards guidance of learning. In this respect Vermunt distinguishes three major approaches to guidance: a 'strong external control' of learning activities (the teacher clarifies subject matter and facilitates, regulates and controls the processing, and minimalizes the need for student to exert thinking energy); a 'loose external control' (teaching is limited to presentation of material and testing) and a 'shared control' (the teacher and the students keep joined control over the learning process).
In combining learning and instruction theories Vermunt gives examples of instructional activities promoting each of the twenty-four different learning activities, both in a 'strong external control'–manner and a 'shared control'–manner.

Construction of an instrument and procedure of the pilot study
This frame of reference was taken to construct an assessment instrument of educational skills. Starting with the twenty-four learning activities, scales were designed, each of which contained three to six items. Since the instrument was to be used in a specific context (students assessing teachers in medical education) not all categories appeared useful distinctions. On the other hand some clear, specific educational activities seemed not to fit in the theoretical model but appeared nevertheless useful to add to the skills to be assessed (technical skills such as fluency of speech, posture during teaching, time–management during lectures, adequately using teaching aids). The instrument eventually contained six 'cognitive processing' scales (most closely related to Vermunts' categories (6) Relating, Structuring, Analyzing, Conretizing, Memorising, Processing critically), four 'motivational' scales (more or less resembling Vermunts' categories combined in Motivating, Concentrating, Appraising and a new category 'Involving students in the subject matter'), four 'metacognitive regulation'–scales (Orienting/Planning, Diagnosing/Evaluating, Repairing, Monitoring/Testing) and finally two added scales: 'Presentation' and 'Use of teaching aids'.
In each of the 'Vermunt'–scales most items concerned teaching behaviour (i.e. 'external control') but each time one item was formulated as 'the teacher stimulates ..(the specific learning activity)', in this way expressing the 'shared control' style of teaching for this activity.
A pilot study was designed to test applicability of the procedure, utility of the items as formulated and reliability of the scales. Ten teachers of the Faculty volunteered for this study. Their classroom performance was assessed by students on voluntary basis. Student–judgements are known to yield reasonably valid measures for teacher assessment (4,1); furthermore, students are the only ones to confirm teacher skills resulting in shared control of (their own) learning activities. A total of 250 students (years 1 through 4) were asked to participate in the pilot study. The assessment instrument contained (A) a total of 72 items, distributed over 16 separate scales, (B) five general questions, indicating (i) the most important items, (ii) a general 'mark' for the teacher, (iii) how much this teacher has stimulated efficient studying, (iv) deep–level learning and (v) investment of time in studying and (C) space for additional remarks on the teacher or the instrument. All

items, formulated as a statement, continued with three questions: 'how much did the teacher this?', 'how well did the teacher this?' and 'is this an important activity for these lessons?'. Classroom performance was assessed in lectures, small group teaching and lab classes.

Results of the pilot study

In total 60% of the students who agreed to participate responded, yielding 306 completed questionnaires.

A standard procedure was applied to select the most useful items for an improved assessment instrument (e.g. constituing homogeneous scales, answerable and realistic items from students' points of view, items related to study behaviour).

The major results can be summarized as follows. Not all types of instruction could be usefully benefit from one single instrument. It appeared necessary to distinguish an instrument for lectures and small group instruction and one for lab classes, containing partly the same, and partly different items.

The 'theory'–classes instrument eventually contained a total of 47 items. Of the 14 theoretical categories, 9 appeared to hold (Relating, Structuring, Memorising, Motivating, Appraising, Involving, Orienting/Planning, Diagnosing/Evaluating/Monitoring/Testing, Repairing). Four added scales could be distinguished: Time management,Eloquential behaviour (both part of 'Presentation'), Use of teaching aids and Stimulating active study behaviour. This last scale represents the 'shared control'. It appeared that adding one 'shared control'–item to all scales unfortunately did not enhance the scales' homogeneity. Consequently some of them were separately combined to form an added scale.

The lab–classes instrument contained 31 items and six scales, all of which belong to the affective and metacognitive categories or presentation. No cognitive processing categories seemed useful.

Of course the psychometric qualities of these new instruments still remain to be tested. This is currently done in a second study.

Discussion

Given space limitations a selection of some general remarks may be made.

Educational competence is not only reflected in classroom performance. Organizing and management of education, testing and preparing or selecting written materials also may have major impact on student behaviour. Assessing educational competence ideally implies also these other qualities. Instruments for these are being prepared too.

The question may be raised whether medical teachers should also possess specific teaching skills, for instance pertaining to demonstration of patients. We believe that these are specific manifestations of general educational skills that may be assessed with a general instrument. Training of teachers, however, will probably be most effective if executed in a specific context (i.e. medical if the teachers in question are clinicians).

External validity of the instrument has not been the major concern in the pilot study. This very important issue – specifically when assessment results are to be reported to individuals – is being addressed in the second part of the project. Validity studies are necessary before developing a policy in which results of assessment may have personnel consequences, such as advice to attend teacher

training. In this respect it will be necessary to realize that quality of education also depends on factors not influenced by individual teachers. Therefore results of teacher assessment will have only a limited – and maybe not detectable – relation to student achievement.

References

Gijselaers, W. H. (1988). Kwaliteit van onderwijs gemeten: Studies naar de betrouwbaarheid, validiteit en bruikbaarheid van studentoordelen. Doctoral thesis. Maastricht.

HOAK (1986). Ministerie van Onderwijs en Wetenschappen. Nota Hoger Onderwijs: Automie en Kwaliteit. Den Haag: Staatsuitgeverij.

Kells, H. R. & Vught, F. A., van (1988). Theoretical and practical aspects of a self–regulation and quality control system for Dutch higher education. Tijdschrift voor Hoger Onderwijs, 6, 1, 15–20.

Marsh, H. W. (1987). Students' evaluations of university teaching: research findings, methodological issues, and directions for future research. International Journal of Educational Research, 11, 3, 257–379.

Verbeek, A. E., Bijtel, J. A. L., van den, & Cate, Th. J. , ten (1992). Instrumentontwikkeling voor de beoordeling van docenten. SCO–report 281. Amsterdam: Stichting Kohnstamm Fonds voor Onderwijsresearch.

Vermunt, J. D. H. M. (1989). The interplay between internal and external regulation of learning, and the design of process–oriented instruction. Paper presented at the 3rd Conference of the European Association of Research on Learning. Madrid.

Wilbrink, B., & Hofstee, W. K. B., (1984). Docentbeoordeling, mogelijkheden en randvoorwaarden. Onderzoek van Onderwijs, 13, 4, 22–25.

Wijffels, H. J. C., (1990). Beoordeling van docenten door studenten als kwaliteitsbewaking van onderwijs. In G. W. H. Heinen, T. H. Joostens, & A. I. Vroeijenstijn (Eds.). Kwaliteitszorg – waarborg voor kwaliteit in het hoger onderwijs. Groningen: COWOG, RUG.

IMPLEMENTING KNOWLEDGE ABOUT LANGUAGE IN THE ENGLISH NATIONAL CURRICULUM

Janet White, National Foundation for Educational Research, United Kingdom

Introduction

LINC (Language in the National Curriculum) is a three-year in-service programme (1989-92) funded by the Department of Education and Science and the local education authorities of England and Wales. Its main aim is to acquaint teachers with the model of language developed from the Report of the Committee of Inquiry into English Language Teaching (Kingman Report, 1988).

The LINC project is in many respects the culmination of a decade of initiatives in the development of English as a curriculum subject; discussion of the project will refer to the nature of contestation surrounding literacy teaching and standards of achievement in English which has been a regular feature of the educational climate in the UK, especially since the introduction of the national curriculum in the late 80's.

The Kingman Report proposed a model for teaching about language that included not only the study of forms of language, eg. sentence structure, but also speaker/listener, reader/writer relationships encompassing the notions of context, audience and purpose; language acquisition and development; and language variation.

The Cox Report (English 5-16, 1989) subsequently made recommendations for the teaching of knowledge about language - providing two justifications for teaching children explicitly about language:

- the positive effect on aspects of their use of language;
- the general value of such knowledge as an important part of their understanding of their social and cultural environment, since language has vital functions in the life of the individual and of society.

Taken together, the Kingman and Cox recommendations have led to the teaching of knowledge about language being a statutory requirement of the National Curriculum for English. Aspects of knowledge about language (KAL) feature in the assessment arrangements for children around age 11 onwards, but as the Programmes of Study make clear, opportunities for pupils to develop this knowledge should be provided from the beginning of schooling.

Method

The LINC project team, comprised of 25 regional coordinators from the local education authorities of England and Wales, have prepared and published materials for teachers to assist them in developing their own and, in turn, their pupils' knowledge about language. The materials for teachers have formed the basis of numerous training programmes based on a cascade principle, whereby regional coordinators worked extensively with Primary advisory teachers and Secondary Heads of English, who then disseminated the work to groups of schools, English departments and individual teachers.

The training has involved teachers in most schools in England and Wales during the period 1989-92 and is based on the following principles: extending pupils' knowledge about language should start from what children know and can do; language should be explored in purposeful situations to encourage reflection and analysis; an understanding of attitudes to language can enable pupils to see through language to the underlying ways in which values and beliefs are communicated in it.

Discussion

The paper will show how these principles have informed the writing of materials for teachers' professional development, and illustrate ways in which teachers' understanding of them has informed classroom practice. The paper will also comment on ways in which the cascade model of in-service training has been variously interpreted in light of local conditions.

TEACHER PROBLEMS AND NEED FOR IN–SERVICE TRAINING

Th. Houtveen, & A.L. van der Vegt, Department of Educational Research and Department of Education (ISOR), University of Utrecht, Utrecht, The Netherlands

Research questions

Many discussions have been held in the last few years concerning the quality of education. The quality of teachers has also been an issue in these discussions. Since the late seventies educational researchers have concerned themselves with the functioning of teachers (cf. Veenman, 1984, 1989; Peters, 1985; Bergen & Van Opdorp, 1989). However, these studies were mostly confined to an inventory of problems of teachers, without giving any insights into the degree to which teachers try to come to grips with their problems, nor into the degree to which teachers would like to receive in–service training in dealing with their problems. We therefore conducted a study which addresses the following questions (Houtveen & Van der Vegt, 1991):

- To what degree do teachers in the first stage of secondary education pay attention to preventing problems?
- To what degree do teachers wish for in–service training in dealing with their problems?

Operationalisation

On the basis of the problem inventories in the literature we made a division into problem areas with which teachers try to cope. The basic assumption in the selection of these areas was the possibility for inservice–training to contribute to the solution of the problems. This led us to the following eleven problem areas:

- working according to a plan;
- handling differences of level between pupils;
- varying teaching methods;
- evaluating pupils performance;
- renewing course contents;
- taking measures to keep order;
- motivating pupils;
- taking a personal interest in pupils;
- having professional relations with colleagues;
- getting support from the organization;

- working collectively on innovations.

These eleven problem areas were categorized into three problem fields, namely 'teaching' (1 – 5), 'teacher–pupil interaction' (6 – 10) and 'working environment' (9 – 11). The first two problem fields concern the classroom level, the third field concerns the school level. This division ties in with the categorisation found in Peters (1985) concerning the levels at which problems can arise for teachers: the micro–, meso– and macro–level. We left out the macro–level since change agents can't contribute to the solution of problems on this level.

A Likert scale was constructed for each of the problem areas.

The first 8 scales measure to what extent teachers take action to prevent problems. We give an example of an item from the scale 'taking measures to keep order': 'I don't start a lesson until it is completely quiet'. The scales which belong to the problem field 'working environment' measure to what extent attention is paid at the school level to the problem areas involved in the perception of the teachers. We give an example of an item from the scale 'getting support from the organization': 'Management and teachers fall in with each other when it comes to order measures'. The research instrument comprises 104 items with six respons modalities (1= completely untrue; 6= completely true).

Respondent group

A total of 616 teachers from the first stage of secondary education took part in the study, comprising 173 beginning teachers (with less than five years of experience), 231 experienced teachers (from six to fifteen years of experience) and 199 senior teachers (with more than fifteen years of experience).

Psychometric quality

In order to investigate the psychometric quality of the research instrument we tested for the homogeneity of the separate scales as well as the construct validity of the three separate problem fields. Two scales ('evaluating pupils performance' and 'having professional relations with colleagues') didn't meet the Cronbach's alpha criterion of at least .70. We therefore removed these scales from the instrument. The other scales turned out to be sufficiently homogeneous to deserve the label reliable.

In order to scrutinize the construct validity of the instrument we designed a theoretical model in which a distinction is made between problems concerning teaching skills, problems concerning the teacher–pupil interaction and problems concerning the working environment (see operationalisation). With the use of LISREL (maximum likelihood) we investigated how far this model fits our data. We found a chi–square coefficient of 29.5 with 24 degrees of freedom and a p–value of .27. This indicates a good fit. These results provide evidence for the construct validity of the instrument.

Results

Working on problem areas

On average teachers turn out to pay at least a fair amount of attention to the problem areas which concern the problem fields of 'teaching' and 'teacher–pupil interaction'. No lower mean score than .52 was found on each of the scales which express these problem fields. The literature mentions the isolation of teachers as

an important problem source. This is not confirmed by our study. The scores on the problem field 'working environment' were high. This means that teachers feel supported by the organization to a high degree. Teachers also indicate that working on innovations is, to a high degree, a matter for the school as a whole. The literature shows clearly that the problems of older teachers are different from those of beginning teachers. In our study however, we found that in the fields of 'teaching' and 'working environment' the differences between beginning teachers, experienced teachers and senior teachers are very small. The amount of experience doesn't seem to effect the degree in which teachers pay attention to these problem fields. By contrast, the field of 'teacher–pupil interaction' shows significant differences between beginning teachers on the one hand and experienced and senior teachers on the other hand, in the degree in which they take disciplinary measures and motivate pupils. This result is in accordance with the fact that in the problem inventories keeping order and motivating pupils are the most frequently mentioned problems by beginning teachers.

Need for in–service training

In particular the field of 'teaching' shows a fairly large percentage of teachers who want in–service training. One in six teachers wish for 'a fair amount' or 'a lot of' in–service training in 'handling differences of level between pupils'; one in five needs in–service training in 'varying teaching methods'; and one in four in 'renewing course content'. The results show further that in the field of 'teacher–pupil interaction' slightly less than 50% of the teachers admits to 'some' need for in–service training in 'motivating pupils' and one–third in 'taking measures to keep order'. As regards the 'working environment' field we see that one–third of the teachers would welcome at least 'some' in–service training in 'working collectively on innovations'.

For practically every problem field beginning teachers show significantly more need for in–service training than their more experienced colleagues.

Finally, the research shows that a fairly large to a large percentage of teachers who wish for in–service training want it to be given by colleagues.

Conclusion and discussion

Teachers from the first stage of secondary education pay, on average, at least a fair amount of attention to the problem areas which occur most frequently in the teaching profession. A fairly large percentage of teachers welcomes in–service training in working on these problems, especially in the field of 'teaching'. The results also show that a considerable percentage of teachers prefers support by colleagues.

Where problem inventories in the past showed particularly clearly that teaching is a demanding profession, the results of the present study offer points of departure for in–service training and further training. Not only with regard to the problem areas, but also with regard to the respective roles of external and internal support. Considering the fact that teachers often prefer in–service training by colleagues, institutions for educational support could concentrate more on training teachers in giving each other in–service training, rather than giving in–service training themselves.

References

Veenman, S. (1984). Perceived problems of beginning teachers. Review of Educational Research, Vol. 54, nr 2, pp. 143–178.

Veenman, S. (1989). Inscholing van beginnende leraren. In S. Janssens, J. J. Peters, & G. Tistaert (red.). Leraar als beroep, pp. 119–132. Amsterdam/Lisse: Swets & Zeitlinger.

Peters, V. (1985). Docenten en hun probleemsituaties. Nijmegen: Katholieke Universiteit.

Bergen, Th. C. M., & Opdorp, C. A. W., van (1989). Probleemsituaties in het onderwijs: Een beoordeling door docenten. Pedagogische Studieën, 66, 327–340.

Houtveen, Th., & Vegt., A. L., van der (1991). Begeleiden van leraren: Onderzoek naar de mogelijkheden van en behoeften aan begeleiding van leraren in de 1e fase van het voortgezet onderwijs. Utrecht: ISOR.

ASSESSMENT OF STARTING COMPETENCES OF STUDENT-TEACHERS

Robert de Jong, Institute of Educational Research (RION), State University Groningen, Groningen, The Netherlands

Introduction

Teacher training colleges for primary education have to be concerned about the quality of their education. Not only because so called 'visiting committees' will report (in 1991/92) about the quality of these colleges but also because colleges should be concerned about the quality of their 'product'. The first criterion for quality should be the competence of students (De Groot, 1986).
Three training colleges were concerned about the starting competence of their students. They asked the Institute of Educational Research (RION) of the university of Groningen to develop an instrument for measuring the starting competences of their students. The instrument should more specifically measure 'generic' competences rather than subject matter expertise, and be useful for teacher-trainers in the assessment of students.
The goal of the research-project was to develop and analyse the quality of the instrument.

Instrument development

Two problems had to be solved in order to be able to develop the instrument. These problems are the specification of 'generic' competences and the kind of method for assessing these competences.

- The 'generic' competences

The job of teacher is labeled as a 'profession'. In contrast to technical jobs teachers can only be judged on using 'the best practice available' (Soar et al., 1983,240). In the USA their is a lot of discussion about the kind of competences (student) teachers need in their job. Process-product studies have yielded findings on the forms of teacher behaviour that most effectively promote student learning (Medley, 1982). These studies are criticized because they ignore one central aspect of classroom life: the subject matter (Shulman, 1986). The knowledge base of teaching has been studied in the Teacher Assessment Project but there are up till now no definite answers to the question what makes teaching effective (Haertel, 1991). There is no simple solution for this problem because effective teaching depends of the context (pupils, subject).
Although the state of knowledge doesn't give definite answers assessment of teachers has to be done.

In order to formulate starting competences of teachers several methods can be used (Haertel, 1991). In this research-project the kind of competences students need in order to be able to start as a teacher was based on earlier research in which competences were derived from a survey among different experts and from literature about process-product studies. De Jong, Tillema & Wolfgram (1985) developed a list of (relevant) competences for primary schoolteachers on the basis of literature and perceptions of teachers, teacher trainers, staff-developers and content-specialists. De Jong, Matthijssen & Van der Wal (1987) performed an analysis of literature for the inspection of higher education about process-product research and formulated a list of effective competences.
The training colleges appraised the competences according to the goals and contents of their curriculum. The result was a list of 82 competences which were devided in 15 task areas.

- The instrument
In the literature a lot of assessment procedures are summarized f.e. observation, ratings, self-assessment, student test scores, portfolios, conventional tests, etc. (Millman & Darling-Hammond, 1990). The choice of the preferred method is dependent of several factors (the kind of teacher, the kind of competences, the kind of assessor, etc.). The instrument should be practical for teacher-trainers and therefore it was decided to use a rating scale. Teacher trainers should assess the competence level of students they teach. Are teacher-trainers who teach different subjects capable in reliably assessing the competences of students?

Method
In the research-project 33 teacher-trainers of 3 colleges participated. They assessed the competence level of 217 students on a 5-point likert scale (1=bad; 5=very good). The students were in their third and fourth (last) year of study. Each teacher-trainer should assess 10 students on each of the 82 competences if possible. Each student was assessed by 2 or 3 teacher-trainers in order to measure inter-rater reliability.

Data-analysis
Many teacher-trainers were not able to assess all the competences and so for many competences there were not enough ratings. For this reason the data-set was reduced to 15 teacher-trainers who assessed 109 students on 41 competences.

Results
The variation of the reliability of the 15 competency scales was between .81 and .98. Further analysis revealed that an important cause for these high reliabilities was the inability of the teacher-trainers to rate differentially the competences within each scale. Many teacher-trainers (dependant of the type of competence between 22% and 66%) gave the same rating to the different competences within one scale.
The inter-rater reliability (between teacher-trainers who assessed the same students) was rather low (spaerman-brown was .07). The correlation between the 'generic' competency ratings and the marks the students received for the subject of the teacher-trainer was rather high (r=.66).

Conclusion and interpretation

The results show that the instrument cannot be used within the colleges for summative purposes. The teacher-trainers assess the students probably more on the basis of the subject achievements than on the basis of what they know of the achievements on the 82 competences.

The results of this research-project can be compared with research about ratings in another context. The inter-rater reliability of experts assessing subsidy proposals are also low (Hofstee, 1983; Scheerens & Beem, 1986).

Suggestions for improving the reliability of the instrument are:

- linking specific competences to specific observors ('experts');
- training of observors;
- stimulating discussions about differences in assessment between teacher trainers.

References

Groot, A. D., de (1986). Begrip van evalueren. Den Haag: VUGA.

Haertel, E. H. (1991). New Forms of teacher assessment. Review of Research in Education, 17, 3-31.

Hofstee, W. B. (1983). Ratings of grant requests in educational research: A psychometric evaluation. TOR, 8, 273-284.

Jong, R., de, Tillema, H. H., & Wolfgram, H. P. (1985). Professionele bekwaamheden en nascholingsbehoeften van leerkrachten basisonderwijs. Eindrapport SVO-project 4125. Groningen: RION.

Jong, R., de., Matthijssen, C., & Wal, G., van der (1987). De brede bekwaamheid binnen de PABO (een tussenstand). Groningen: RION.

Jong, R., de (1991). Kwaliteit moet blijken. De beproeving van een instrument voor de beoordeling van de brede bekwaamheid van PABO-studenten. Groningen: RION.

Millman, J., & Darling-Hammond, L. (ed.). (1990). The new handbook of teacher evaluation. London: Sage.

Scheerens, J., & Beem, A. L. (1986). Ratings of requests for grants in Educational Research revisited. TOR,11,2, 67-82.

Soar, R. S., Medley, D.M., & Coker, H. (1983). Teacher evaluation: A critique of currently used methods. Phi Delta Kappan (dec), 239-246.

A STUDY ON THE ADMISSION STANDARDS FOR TEACHER TRAINING PROGRAMS

Fersun Paykoç, & Ahmet Ok, Department of Educational Sciences, Middle East Technical University, Ankara, Turkey

Improving the quality of education is largely attributed to the improvement of teacher quality. Careful selection of students for teacher training programs is one of the attempts put forth to raise the quality of teachers and education (Sandefur, 1984; Felstritzen, 1984). It has also been revealed through theory and research in vocational guidance and curriculum development that matching the characteristics required in a professional program and the individual's entry characteristics has positive effects on one's performance. (Bloom, 1976; London, 1973)
According to Awender and Harte (1986) admission standards for students are one of the dimensions attributed for increasing the quality of teacher education programs. Cruicksbank and Cruz (1989) also emphasized the establishment of higher standards for the admission of students into teacher education programs. Parallel to the emphasis observed in literature, importance of entry characteristics in the success of students and in the effectiveness of teacher training programs have also been considered vital in Turkey. As the university entrance examination is the only measure for the admission of students to teacher training programs, other measures are searched for. The main purpose of this study was to examine certain admission standards in the students already admitted into teacher training programs. This study is the second part of a larger one which first aimed to determine the standards through using the Delphi technique with the consensus established by teachers, academicians, principals and parents.

Method

The subjects n=113) consisted of Science Education (n=89) and Foreign Language Education (n=24) freshmen first term students in the Faculty of education at M.E.T.U. Ankara Turkey.The data related to the standards was collected in the first part of the study through the Delphi questionnaire administered to 37 teachers, 41 academicians, 33 principals and 27 parents. The data for the second part was collected from a set of already developed attitude scales, an information sheet and student files.
The first scale was intended to measure attitudes toward teaching as a profession. (Askar & Erden, 1987) The other two scales were used to determine subject related attitudes of the students of mathematics, chemistry, biology and physics teaching programs. Attitudes toward mathematics was measured through a scale

developed by Aiken. (Aiken, 1979). The other subject related attitude scales in chemistry, biology and physics were also reproduced from Aiken (1979).
The data about student's preference ranking of the program in which they were placed by the Student Selection and Placement Centre in Turkey and the reason for preferring the teacher training program were collected through the information sheet. The data on student's background information (high school and university entrance) was collected from the student files at the Registrar's Office. The university entrance examination involves two stages; the student selection examination (SSE) includes a verbal and a quantitative sub-test. The student placement examination (SPE) includes a test for natural. Sciences, a test for mathematics, a test for Turkish language and literature, a test for social sciences and a test for foreign language (ÖSYM, 1988). The data analysis was carried out through the sub-programs of SPSS (Nie, et al., 1975).

Results

The following admission standards were determined through the Delphi technique: in the first part of this study:

- Ability to understand written expressions and concepts (x=5.58, s=0.56, Agreement: 100 %)
- Interest in teaching (x=5.50, s=0.68, Agreement : 100 %);
- Ability to express oneself (x=5.49, s=0.65, Agreement : 98 %);
- Attitudes toward teaching as a profession (x=5.36, s=0.93, Agreement : 96 %);
- Knowledge of vocabulary (x=5.21, z=0.77, Agreement: 97 %);
- Subject related attitudes (x=5.17, s=0.90, Agreement: 95 %);
- The prerequisite knowledge and skills needed to be successful in a teacher training program (x=5.07, s=0.95, Agreement: 95 %);
- Ability to establish quantitative relationship (x=4.44, s=1.07, Agreement: 81 %);
- Achievement level at high school (x=3.87, s=1.31, Agreement: 70 %).

In relation to the admission standards; "the prerequisite knowledge and skills needed", "Ability to establish quantitative relationships", "Ability to understand written expressions and concepts", "Ability to express oneself" and "Knowledge of vocabulary" the related mean scores of the Student Selection Examination (SSE) and Student Placement Examination (SPE) were used for analysis.
The mean scores of SSE of students already admitted into teacher training programs indicated that the students had about 25 to 50 points higher scores when compared to the cut-off score required for being invited to the second stage of UEE. The SPE mean scores were not comparable among the different teaching groups. They were examined with in their own range. Although the means of mathematics, chemistry and physics teaching students were different, they were admitted on the basis of the same composite scores.
Similarly the means of correct answers differed for the Mathematics sub-test of UEE. None of the science groups had more than 50 % correct answers except for the mathematics teaching students. For the sub-test on natural sciences none of the groups had more than 50 % correct answers. English teaching students had more than 50 % correct answers on the social sciences sub test and had 68 correct answers out of 75 on the English sub-test of UEE.
When the high school achievement was compared on the basis of diploma grades for the five groups they were mainly similar except for the mathematics teaching

group being higher than the others. The high school grade point average was included in the UEE and about 10 % of the total respective weights was related to this measure. In relation to the representation of the verbal/language abilities, the related subtests were examined and it was observed that more than 50 % of the questions in the Turkish language subtests were correctly answered by the students.

Concerning the subject related attitudes, the findings indicated that they tended to be positive. In relation to the findings on attitudes toward teaching as a profession, it was observed that the attitude mean scores were not very high. The means were concentrated around "undecided" or "close to agree". As a measure of the interest in teaching, the ranking of the programs by the students indicated that 95 % of English teaching students and 54 % of mathematics teaching students ranked their program among the first five choices in the UEE. Concerning the reasons of selecting teaching, it was found that 53.6 % of mathematics, 33.3 % of physics, 27.3 % of chemistry and 20.8 % of English teaching students preferred the program to become teachers. As a conclusion, it was generally observed that attitudes toward teaching as a profession and attitudes toward subject matter, were not included and the others were partially observed. In the present higher education admission system.

Conclusions

The existence of the standards were examined only on the basis of the available data. The measures related to the UEE were not all comparable for all the teaching groups. Already developed scales were able to cover some of the standards partially. The high scores observed in some of the standards were positive clues for strengthening the validity of the standards determined through the Delphi technique.

As this was a pilot study, the authors preferred not to reach any generalizations, but to make use of the findings for initiating the development of new measures for the high priority admission standards and to use such standards-especially affective ones-for admitting students in to teacher training programs. It is believed that this will contribute to the improvement of teacher quality in Turkey, and initiate the development of instruments relevant for the determined standards.

References

Aiken, L. R. (1979). Attitudes toward mathematics and science in Iranian middle schools. School Science and Mathematics, 76, 229-234.

Askar, P. ve Erden, M. (1987). Ögretmenlik meslegine yönelik tutum ölçegi. Çagdas Egitim, 12(121), 8-11.

Awender, M. A., & Harte, A. E. (1986A). Teacher education: What is wrong with the current practice. Education Canada, 26(1), 25-27.

Bloom, B. S. (1976). Human Characteristics and school Learning. New York: McGraw-Hill.

Cruicksbank, R. D., & Cruz, J., Jr. (1989A). Trends in teacher preparation. Journal of Teacher Education, 40, (1), 49-50

Feistritzer, C. E. (1984). The making of a teacher education and certification. Washington: National Centre for Education Information.

London, H. H. (1973). Principles and Techniques of Vocational Guidance. Ohio: Charles E. Merril.

Nie, N. H., Hull, C. H., Steinbrenner, K., & Bent, D. H. (1975). Statistical packages for social sciences. (2nd ed.) New York: McGrow Hill.

ÖSYM. (1988). The system of student selection and placement in higher education institutions in Turkiye. Ankara, Meteksan.

Sandefur, J. T. (1984). Standards for admission to teacher education programs. An issue paper presented at for the Minnesota Higher Education coordination Board.

CONTROL AND DISCIPLINE: A PROBLEM FOR BEGINNING TEACHERS?

Warwick Taylor, Janet Draper, & Helen Fraser, Moray House Institute of Education, Heriot-Watt University, Edinburgh, Scotland, United Kingdom

Introduction and Method

The extensive literature on beginning teachers puts control and discipline high on the list of problems perceived to be of significance to such teachers as they start their teaching careers. A review of the literature also indicates that evidence from national samples is sparse, as are studies employing a developmental perspective. In Veenman's (1984) seminal review of the literature on the problems of beginning teachers, of the 83 studies reported on from 1960 onwards, only 5, (Taylor & Dale, 1971; Hooymayers, et al 1976; Frech & Reichwein, 1977; Tisher, Fyfield & Taylor, 1979; Bergen et al 1983;) had national samples as the main data source.

In Scotland this research team has just completed in December 1991 a national study of beginning teachers covering the first two years of teaching (known as the probationary period) and has begun to provide data from a national sample which allows, amongst other findings, substantive comments to be made from evidence which is both wide ranging and developmental in relation to the experiences of beginning teachers in Scotland.

The Scottish primary and secondary school system, which is culturally, legally and administratively separate from the English school system, employs teachers who have completed 4 year degrees in teacher education or one-year post graduate courses. After entering employment in a Regional Authority (local government) such teachers are provisionally registered as teachers on probation with the General Teaching Council for Scotland for two years, that is 380 teaching days. The General Teaching Council for Scotland is a professional body, independent of both the employers and government, and the decision on final registration lies therefore with the profession through the headteachers of the schools in which the probationers are employed. The period of probation is conceived of by the Council as one of continued training and of early professional development.

This paper will concentrate on issues of control and discipline only, although the Study of Probationers project, funded by the Scottish Office Education Department, collected and analysed over a three-year period extensive data in the areas of initial training; recruitment; selection and induction; support and assessment.

843

Interview and questionnaire evidence was drawn from a large sample of probationers in primary and secondary schools, from their headteachers, from the employing authorities and from the initial training institutions as well as the interim and final assessment reports sent to the General Teaching Council.

Probationers and Discipline - Results and Discussion

For the purpose of the conference the data were interrogated for indications of the extent of control and discipline problems experienced by beginning teachers. Although previous research, as reported by Veenman (1984), notes that control and discipline problems are endemic in beginning teacher experiences, the Scottish data do *not* support this. While some 60% of primary probationers report it as a problem at some point in probation, only 10% of primary report it as a continuing problem. No clear pattern is discernible in comparing four year and one year trained primary probationers. While 80% of secondary probationers report discipline as a problem at some stage, on 17% mention it as a continuing problem. These data should be considered in the light of (a) the frequency of discipline problems encountered by experienced teachers and (b) other problems experienced by probationers. In ranking reported problems by frequency of mention, control and discipline features quite significantly in first year but drops significantly by second year. It would seem that when Scottish probationers experienced discipline problems (and a fair number did not) these were generally not continuous.

Headteachers with probationers were asked to identify the three most important things they looked for when assessing probationers. They ranked control and discipline 7th in primary and 6th in secondary. Other criteria were more important, especially class organisation and management and relationships with pupils. And indeed the most commonly reported means of assessment (inspecting forward plans, observing probationer interaction with staff) do seem unlikely to yield systematic data on discipline.

However, the actual *practice* of probationer assessment for final registration would seem to elevate the significance of control and discipline. Control and discipline is an area on which, at the end of the probationary period, headteachers in both primary and secondary sectors, are reluctant to tone down critical comment. The headteachers also reported that school characteristics were taken into account in assessing probationers.

An emphasis on successful discipline is no barrier to professional registration for those who do not experience it as a problem. Its significance in assessment appears to be crucial however, for the small number who do experience it as a continuing problem but its importance as a criterion for success may not have been recognised, nor highlighted in induction.

The data reported were derived from four questionnaires completed by probationers at staged intervals over a two year period and from two questionnaires to their headteachers. The study by D'Arcy (1988) reported a changing pattern of problem areas for beginning teachers which, in general are reflected in these data. This Scottish study however has identified a fluidity and complexity of expressed need of beginning teachers which suggests the fruitfulness of future research targetted at a more rigourous examination of individual need and patterns of professional development of beginning teachers.

Such a European wide co-operative research project awaits the attention of the educational research community.

References

Bergen, Th. C. M., Peters, V. A. M., & Gerris, J. R. M. (1983). Een exploratieve studie naar de probleemsituaties van ervaren en beginnende docenten tijdens hun beroepsuitoefening. In S. A. M. Veenman, & H. A. M. Coonen (Eds.). Onderwijs en Opleiding. Lisse: Swets & Zeitlinger.

D'Arcy, J .(1985). Into Teaching. Belfast: Northern Ireland Council for Educational Research.

Frech, H. W., & Reichwein, R. (1977). Der vergessene Teil der Lehrerbildung: Institutionelle Bedingungen und inhaltliche Tendenzen im Referendariat der Gymnasiallehrer. Stuttgart: Klett-Cotta.

Hooymayers, H. P., Créton, H. A., Genderen, D, van, Hendriks, A. F, Holvast, A. J., & Schröder, J. F. (1976). Beginnende natuurkundeleraren in het HAVO en het VWO. Utrecht: Vakgroep Natuurkunde-Didaktiek R U Utrecht.

Taylor, J. K., & Dale, I. R. (1971). A survey of teachers in their first year of service. Bristol: University of Bristol, Institute of Education.

Tisher, K. P., Fyfield, J .A., & Taylor, S. M. (1979). Beginning to teach: The induction of beginning teachers in Australia (Vols 1, 2). (ERDC Reports No 15, 20). Canberra Australian Government Publishing Service.

Veenman, S. (1984). Perceived problems of beginning teachers. Review of Educational Research, 54, 143 - 178.

IMPROVING THE TEACHING OF BEGINNING SCIENCE TEACHERS: A STUDY OF EXPERT AND NOVICE CLASSROOM PRACTICES

A. Cachapuz, University of Aveiro; M. Jorge, University of Trás-os-Montes, Portugal

"A wise man learns from experience and an even wiser man from the experiences of others"
Plato

Introduction

The recent reform of the whole of the Portuguese school science curriculum highlights the need of skilled science teachers who will be responsible for teaching science for understanding. This is a major challenge in particular for novice teachers for whom covering a given content is often equated with pupils' learning.

A possible strategy to stimulate innovative efforts of beginning science teachers is to gain a better knowledge of critical differences between expert and novice teacher practices. The basic assumption is that the comparison between expert and novice classroom practices may provide useful insights in order to design adequate training courses helping beginning teachers to improve their teaching. The above rational reflects a development of traditional research on exemplary teaching (Tobin & Frazer, 1990; Tobin & Garnett, 1988). Making apparent key differences between expert and novice classroom practices is more informative than simply describing exemplary practices. In particular, knowing what has to be changed may make it easier for the novice teacher the indispensable process of personal reflection about their own teaching.

The purpose of this study is to describe the findings of a case study designed to investigate critical differences in the classroom practices of two elementary science teachers, an expert teacher and a beginning teacher. The study is part of a more general research project involving the authors and aiming to promote teaching for understanding of Portuguese science teachers.

Method

The two teachers were carefully selected. The expert teacher (aged 30) was a teaching training supervisor appointed by the Central Education Office, with 10 years of teaching experience and considered by her colleagues as a *good teacher* as well as a *good supervisor*. The novice teacher (aged 23) was teaching for the

first time (having obtained average grades during her pre-service training). Both teachers had followed their pre-service training (5 years), in the same University. They were teaching elementary science (5^{th} and 6^{th} grades, 11 and 12 years old pupils, respectively) in the same school located in a medium size town in the north of Portugal. The classes were mixed ability classes with a majority of pupils coming from families with medium and high socio-economic background.

An ethnographic research methodology was adopted with a focus on the obtention of qualitative data referring to three broad domains: subject matter knowledge, pedagogical knowledge and content pedagogical knowledge (as defined by Shulman, 1987). Three complementary methods were used to gather the data:

- Classroom observation of four lessons (each teacher) over a two week period. Narrative records taken by the researchers during the lessons helped to clarify specific issues (e. g. management strategies);
- In depth semi-structured interviews as well as informal interviews (after each lesson) were conducted to clarify reasons behind classroom practices and expert and novice teacher perspectives about science teaching and learning;
- Analysis of curriculum materials used, in particular worksheets distributed to pupils.

All lessons and formal interviews were recorded and subsequently translated into written protocols to make content analysis easier (nearly 110 A_4 pages).

Evidences obtained from those multiple sources were constantly cross-compared (within and between comparisons) to ensure validity. Inconsistent evidences were ignored. Evidences used to infer these features from segments of data were validated by two independent judges, experts in science education (94% and 98% of agreement).

Results

The results suggest five key features discriminating between the teaching of the two teachers (Table 1).

Criterial features identified	Examples of successful practices (expert position)
Science subject matter knowledge	no errors or ambiguous explanations
Active instructional strategies	use of interpretative function of language (emphasis on meaning rather than on definitions)
Responsive environment	suitable feedback to monitor pupils learning
Management strategies	small groups interactions to explore collaborative efforts
Hollistic view of the science curriculum	bridging different topics of the curriculum

Table 1
Main features discriminating between expert and novice practices

The most striking difference concerned the subject matter knowledge. For example, when the novice teacher was discussing criterial differences between living and non-living beings, she emphasised the ideas of "growth" and

"movement" instead of the idea of "assimilation". This probably explains why some pupils provided (in their homeworks) as examples of living beings some rocks (e.g. granite). The lack of confidence in her scientific knowledge also influenced some of the instructional strategies adopted. In fact, whereas the expert teacher explored the language to help pupils to construct their own meanings (interpretative dimension) the novice teacher was mainly concerned in providing definitions. For example, when introducing the concept of "cell" the fundamental idea of "cell" as the basic unit of living beings was probably lost because of the emphasis on definitional aspects. She was also not very well succeeded when exploring analogies (e.g. cell membrane like a raincoat). The expert teacher was able to use different instructional strategies (including a small investigation) giving pupils the opportunity to use their language in a speculative way (e.g. when discussing the origins of granitic layers, *... may be it was an old volcano and then may be the earth crust ... hum ... closed up*) and simultaneously promoting a dynamic perspective of scientific knowledge (e.g. views about earth evolution). She successfully explored collaborative efforts among the pupils avoiding individualistic and competitive classroom environments; there were few evidences of target students, specially for interpretative and evaluative questions (higher levels of thinking). The expert was able to create a structured learning environment defining goals, making the logic of the activities and discussions explicit and specifying what pupils were expected to do. This contrasted sharply with the novice teacher who often seemed to promote discussion for the sake of it without a clear purpose. Interview data suggested that the novice teacher was not even aware of the fuzzy ways she organized her teaching.

Conclusions

Overall, the results suggest that the expert teacher was more concerned with pupils' understanding whereas the novice teacher was more concerned with "survival" strategies. The expert teacher succeeded in her efforts because she was able to integrate subject matter knowledge and pedagogical knowledge in such a way as to make science content more understandable to the pupils. The development of such a complex process needs time to take place. However, time is not the key variable. Becoming a skilled teacher is not a linear process and previous research indicates that the shift from initial to more mature stages of teaching may be facilitated by the use of systematic and intentional reflexion about teaching (Schön, 1987). Reflexion should therefore be used as a strategic tool to bring about changes in the education of teachers.

Findings of the type reported in this case study may provide useful cues to design in-service courses to improve science teaching. The focus of these courses should be on the critical analysis of the identified successful teaching aspects and on the conditions under which it took place. This is in line with proposals put forward by Penick and Yager (1983):

> *What we wanted were case studies of excellence in science education that would guide and provide support for innovative efforts. By heeding case studies ... we thought that much of the trial and error associated with educational innovation could be eliminated and school programs could begin building on what was known to work* (p. 621).

Thus we may help to bridge educational research to teacher education.

References

Penick, J. E., & Yager, R. E. (1983). The search for excellence in science education. Phi Delta Kappan, 64, 621-623.

Schulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57, 1-22.

Schön, D. (1987). Educating the reflective practioner. San Francisco, Jossey-Bass.

Tobin, K., & Garnett, P. (1988). Exemplary practice in science classrooms. Science Education, 72(2), 197-208.

Tobin, K., & Frazer, B. (1990). What does it mean to be an exemplary science teacher? Journal of Research in Science Teaching, 27(1), 3-25.

849

INTRODUCING THE REALITY SHOCK INTO PRESERVICE TEACHER EDUCATION

C.P. Koetsier, J.S. ten Brinke, Th. Wubbels, & C. van Driel, Institute of Education and Centre for Science and Mathematics Education (IVLOS), University of Utrecht, Utrecht, The Netherlands

Introduction

Some decades ago, teaching in secondary education was learnt mainly by just doing it. This trial and error strategy, however, was not thought to produce sufficiently competent teachers. For that reason, teacher education (TE) programmes were developed that included student teaching periods in which the learning process could be monitored. Now a new problem materialized, namely that beginning teachers were confronted with a so-called reality shock (Müller-Fohrbrodt, Cloetta & Dann, 1978; Veenman, 1984). One of the aspects of this shock is the gap between ideal and reality. Teacher education was blamed for not adequately preparing student teachers (STs) for the reality of the teaching profession.

One of the possible solutions to this problem, is to try to introduce the reality shock into the TE programme. By doing that, the 'handling' of such a shock by the ST can be supervised. In order to ensure the appearing of (a) reality shock(s) within the TE programme, the practical component of the programme must be arranged in such a way that it shows a maximum similarity to the work situation of the beginning teacher. In practice this means that this part of the practical component should be characterized by insecurity, complexity, need for independent and responsible action, and a substantial work load.

In order to investigate whether it is possible to create a student teaching period satisfying the conditions just mentioned, we have experimented with a new student teaching element (Koetsier, 1991). These experiments have been executed within the context of the one year postgraduate TE programme, with a practical component of 850 hours, which prepares teachers for (higher-level) secondary education. The practical component consists of two elements. The latter, new, element is the Individual and Independent Final Teaching Period (IFTP), which is introduced after the first element, the Triad Student Teaching Period. The Triad Period is a comparatively secure teaching period because it is done by a group of three students together and is intensively supervised. The Triad Period is followed by a reorientation period, intended for reflection, the deepening of insights, and

preparation for the IFTP. The IFTP is a continuous three-months student teaching period and takes place in the last part of the one-year postgraduate TE programme. The ST, who is presented to the pupils as a qualified teacher, teaches all lessons - 10-12 lessons a week - in a number of classes. The ST is supervised 'from a distance' by a cooperating teacher (CT) and a university supervisor. 'From a distance' means that these persons do not attend the STs lessons.
This paper deals with the following three questions: 'What are in the opinion of the respondents important characteristics of an IFTP?', 'Which of these characteristics helps to make the IFTP a realistic teaching situation and which to ensure a favourable learning situation?', and 'How do respondents assess the STs' learning outcomes, especially as to the question in what respects the STs have learnt more or different than during the previous Triad Period?'

Methods
Data were collected during pilot studies in six schools, by means of extensive interviews with those involved in the IFTP experiments (STs, CTs, university supervisors, and headmasters: 24 respondents in all), as well as participatory observation and document analysis. Several strategies were used to ensure validity and reliability, including various forms of triangulation, replication, peer debriefing, member checks, back coupling, and cross checks (Koetsier, 1991) with the aid of procedures developed by a number of authors (including Glaser & Strauss, 1967; Mann, 1970; Denzin, 1978; Smith, 1978; Guba, 1978, 1981; Spradley, 1979, 1980; Patton, 1980; Hycner, 1985).

Outcomes
The IFTP appears to be a kind of student teaching period that shows a strong similarity to the work situation of a beginning teacher and confronts the ST with the normal problems of beginning teachers. The respondents agree that the reality shock seems indeed to have been introduced into the preservice TE programme. The respondents unanimously agree that, for being as realistic as possible, an IFTP should have the following seven characteristics. An application procedure should be used for admitting an ST to the IFTP in a particular school. The IFTP should take place at the beginning of the school year, at last up to the first school report. It should take place in another school than the Triad Period. The ST should be introduced to the pupils as a qualified teacher. A realistic work load should be created: 10-12 periods a week including all other duties of a teacher that are involved. Supervision should not include attending of lessons.
In the respondents' opinion, the STs, in comparison with the Triad Student Teaching Period, acquire more organizational, relational (communicative), educational and exploratory/reflective skills and skills to bring about changes in their functioning and their teaching. This means, for instance, long-term planning, remedying mistakes and overseeing several classroom situations at the same time. In the Triad Period, more gains in the domain of pedagogical skills are made.

Conclusion and discussion
The IFTP-experiments provided answers to the three research questions mentioned before. As to the third question (about learning effects), the answers have the form of indications.

Although during the collecting and processing of the data careful validation procedures were employed, the outcomes must be seen in a proper perspective: the experiments involved a fairly small number of individuals. The conclusions must therefore be seen as grounded hypotheses rather than as firm assertions.

References

Denzin, N. K. (1978). The research act. Chicago: Aldine.

Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory. Chicago: Aldine.

Guba, E. G. (1978). Toward a methodology of naturalistic inquiry in educational evaluation. Los Angeles: University of California, Centre for the study of evaluation.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. Educational Communication and Technology Journal, 29, 75-91.

Hycner, R. H. (1985). Some guide-lines for phenomenological analysis of interviewdata. Human Studies, 8, 279-303.

Koetsier, C. P. (1991). *Een brug tussen opleiding en praktijk. Een studie over de individuele eindstage als brug tussen lerarenopleiding en praktijk*. [Bridging the gap between teacher education and teacher practice.] Dissertatie. Utrecht: WCC.

Mann, F. C. (1970). Human relation skills in social research. In Filstead.

Müller-Fohrbrodt, G., Cloetta, B., & Dann, H. D. (1978). Der Praxisschock bei jungen Lehrern. Formen - Ursachen - Folgerungen: Eine zusammenfassende Bewertung der theoretischen und empirischen Erkentnisse. Stuttgart: Klett-Cotta.

Patton, M. Q. (1980). Qualitative Evaluation Methods. London: Sage.

Smith, L. M. (1978). An envolving logic of participant observation; educational ethnography and other case studies. In S. Shulman (Ed.). Review of research in education, 6 , 316-377. Peackock: Itaca.

Spradley, J. P. (1979). The ethnographic interview. New York: Holt Rinehart Winston.

Spradley, J. P. (1980). Participant observation. New York: Holt Rinehart Winston.

Veenman, S. A. M. (1984). Perceived problems of beginning teachers. Review of Educational Research, 54, 143-178.

A TYPOLOGY OF STUDENT TEACHERS' COPING REACTIONS

Wilfried Admiraal, Fred Korthagen, & Theo Wubbels, Institute of Education and Centre for Science and Mathematics Education (IVLOS), University of Utrecht, Utrecht, The Netherlands

In the paper a study will be presented in which coping behaviours have been clustered into coping reactions that describe combinations and sequential patterns of coping behaviours in response to stressful classroom events. This study forms a part of a research project, in which types of coping reactions will be related to cognitive appraisal of the events, characteristics of the teachers and stressful events, and eventually to the quality of classroom teaching and job satisfaction in order to infer indications for enhancement of adequate types of coping reactions in pre-service teacher education.

Theoretical framework

In our research, coping behaviours are defined as 'behavioural efforts of a student teacher during a lesson to manage classroom events that are appraised as taxing or exceeding the resources of the student teacher' (cf. Lazarus & Folkman 1984). One's coping reaction on a stressful event describes a combination of coping behaviours, that are used to manage a stressful classroom event. In a model of reactions on stressful job situations (c.f. Rusbult, Farrell, Rogers & Mainous III 1988), dissatisfaction in close relationships (c.f. Rusbult, Zembrodt & Gunn 1982), and on minor or major hassles (c.f. Hobfoll 1988), four categories of coping behaviours are distinguished based on two dimensions. The first dimension, which can be described as an approach dimension, concerns to what extent one's behaviours are oriented towards the stressful event or away from it. This dimension is analogous to a common classification into problem-focused and emotion-focused coping behaviours. The second dimension can be described as an activity dimension, ranging from action to inaction. In a combination of the two dimensions, four categories of coping behaviours can be defined: 'direct attack of the stressful event', 'passive response oriented towards the stressful event', 'flight response', and 'neglect'.

The present study is designed to develop a typology of student teachers' coping reactions in response to stressful classroom events, based on the four categories of coping behaviours described above.

853

Method

Data gathering procedure

The participants in this study were 27 student teachers in the teacher education programme of Utrecht University, who had little or no teaching experience. One lesson of each teacher, that was part of the student teacher assignment, was video-taped. After the lesson, an interview was administered, based on a stimulated recall procedure. Parts of the lesson that were appraised as stressful by the student teacher were replayed and the student teacher was asked to recall his or her feelings, cognitions, and actions. Records of the interview session have been transcribed.

Categories of coping behaviours

Student teachers' coping reactions have been mapped by observing excerpts of the recorded lesson that were appraised as stressful. Each coping reaction consists of one or a series of coping behaviours in response to one stressful classroom event. A coding scheme was used, in which four categories of coping behaviours are distinguished: 'direct attack the situation', 'passive responses oriented towards the situation', 'flight responses', and 'neglect'. For each reaction on a stressful classroom event (from a total of 284) scoring units have been determined by event sampling, based on a change in the nature of teaching behaviours. Each reaction has been divided into a series of scoring units (ranging form 1 to 72) and results in an identical number codes of coping behaviours.

Analyses

The unit of analysis is the reaction on each stressful classroom event. Reliability of the scoring procedure was established and validity of the instrument of coping behaviours was checked with factor analysis. In order to identify a typology of coping reactions a cluster analysis technique was used. As input served correlations between frequencies of the four categories of coping behaviours and of correlations between the duration in seconds of the four categories of coping behaviours in a reaction. Besides cluster analysis, we used behaviour sequential analyses in order to determine sequential patterns of coping behaviours.

Results

Reliability and validity of the categories of coping behaviours

Two observers scores a third of the total of stressful classroom events. The categories formed a reliable measurement of coping behaviours (Cohen's $\kappa = .90$). The event sampling based on a change in nature of teaching behaviours showed to be a reliable procedure as well (Cohen's $\kappa = .93$).

Factor analysis of coping behaviours based on time-sampling (5 seconds) revealed evidence for the proposed four categories. For each five seconds within a reaction the four categories were scored on Likert-type scales, ranging from 1(not at all) to 5(very much). After examining the sequences of coping behaviours we noticed sequential patterns of coping behaviours within a reaction.

Results concerning the convergent validity, consisting of correlations between observed coping reactions and self-reported coping reactions, measured in the stimulated recall interviews will be presented.

Typology of coping reactions of student teachers

Results concerning the typology of student teachers' coping reactions will be presented. A typology of coping reactions will be described in three ways. Firstly, we will create a typology of coping reactions based on the percentages of coping behaviours in response to a stressful classroom event. Secondly, we will use cluster analysis to identify types of coping reactions, based on correlations between frequencies and the duration in seconds of the four coping behaviours distinguished. Thirdly, we will use sequential analysis techniques to identify types of coping reactions, based on the sequential patterns of coping behaviours in response to a stressful event. The usefulness for future research on stress and coping reactions of these three ways to identify a typology of coping reactions will be discussed.

References

Hobfoll, S. E. (1988). The Ecology of Stress. New York: Hemisphere Publishing Corporation.

Lazarus, R. S., & Folkman, S. (1984). Stress, Appraisal, and Coping. New York: Springer Publishing Company.

Rusbult, C. E., Farrell, D., Rogers, G., & Mainous III, A. G. (1988). Impact of exchange variables on exit, voice, loyalty, and neglect: An integrative model of responses to declining job satisfaction. Academy of Management Journal, 32, 599-627.

Rusbult, C. E., Zembrodt, I. M., & Gunn, L. K. (1982). Exit, voice, loyalty, and neglect: responses to dissatisfaction in romantic involvements. Journal of Personality and Social Psychology, 12, 1230-1242.

EXPERT AND NOVICE HISTORY TEACHING: EIGHT DIFFERENCES BETWEEN EXPERIENCED AND BEGINNING HISTORY TEACHERS

Eero Ropo, University of Tampere, Finland

Research on expertise in teaching has accumulated rapidly during the 1980's. One of the basic questions in this research has been what kind of differences there are between the experienced and novice teachers. The focus has been in the knowledge differences, but the studies have also described differences in the teaching behaviour. The purpose of the present study is to describe the differences between the conceptions and actual teaching practices of the selected expert and novice history teachers.

The data were collected by interviews and classroom observations. Two to four ordinary history lessons of each subject were observed. The lessons and interviews were tape-recorded and the analyses were based on the transcribed recordings. Results of the analyses were cross-checked independently by two researchers.

The subjects (n=4) were two novices and two experts. As the study served as a pilot for a larger investigation and was aimed to help in finding the directions of teachers' development, a small number of subjects was considered sufficient. No generalizations will be made on the basis of the results.

The experts were selected for the study on the basis of their reputation among colleagues. Both of them had more than 15 years' experience as history teachers. The novices were newcomers, having only one or two years' teaching experience after teacher education.

Results of the study showed that the subjects differed from each other in eight respects. First, there were structural differences in the dispositions of the delivered lessons. Experts' lessons were organized around one or two main themes whereas the novices seemed to have several separated parts in their lessons. Second, there seemed to be qualitative differences in the teachers' conceptions of teaching history. For instance, the experts expressed more opinions of integrating the objectives of history teaching into the overall goals of education. They also seemed to have a better conception of the important and meaningful aspects of history to be learned by the students. Third, planning of the lessons seemed to have a different role for the experienced teachers than for the novices. Basically there were two aspects in planning.The expert teachers did their long-term planning for the school term or the school year. However, those plans seemed to be flexible and based on the main goals of instruction. The novices usually told more about planning of individual lessons. For them planning seemed to consist of

collecting materials to be used in teaching. The novices also planned their demonstrations of different aspects of the history curriculum.
Fourth, the timing of instructional actions was different for the experts and the novices. Typical for the expert teachers was acting as if they had a clock inside them. It meant that the expert teachers noticed the flow of time better than the novices. For instance, they seemed to close discussions just at the correct moment to give a summary before the end of the lesson. Fifth, the expert teachers' conceptions of the need of an additional instructional material was different from that of the novices. Sixth, the teachers differed in respect to the automaticity of their teaching routines. Seventh, there was a difference in the questioning and the quality of questions presented. Eighth, the language of the teachers during the lessons differed between the experienced and novice teachers. During the interviews all the subjects used their native accents. However, during the lessons the experts used common everyday language whereas the novices spoke similar way both during the interview and the lessons. There were also differences in the correctness of the logic of the spoken language and in the sentence structures. The sentences of the novices were more often incomplete than those of the experts.
The differences between the expert and novice teachers include some which has not found before. All the differences seem to show that the novice teachers are in the process of developing knowledge structures for effective teaching. The result of the developing knowledge structures is that the novices have to pay more attention to their own behaviour, whereas the experts can concentrate more on, for instance, students and subject matter. Implications of the study for our knowledge of teacher development, teacher education and research on expertise will be discussed more thoroughly in the paper.

857

THE STATE AND THE TEACHERS: PROBLEMS IN TEACHER SUPPLY, RETENTION AND MORALE IN ENGLAND AND WALES

Gerald Grace, School of Education, University of Durham, Durham, United Kingdom

Abstract

Supply, retention and morale problems related to teaching are in themselves the surface manifestations of deeper historical, structural and ideological contradictions within state policy in education. Such issues cannot possibly be explained without reference to the radical changes in state policy in education which have transformed both the terms and conditions of service for teachers and the nature of the educational process in England and Wales. This paper will attempt to place problems of teacher supply, retention and morale in a wider analytical context by outlining recent historical trends in teacher-state relations which are relevant to contemporary policy issues. It will then proceed to examine the consequences of these changing social and power relations for teacher supply, retention and morale problems. Finally it will attempt a critical reflection upon the options available for a contemporary education policy which is informed by the insights of critical scholarship rather than by the decontextualized analysis of policy science.

858

THE METACOGNITIVE TEACHER

Maria Cardelle-Elawar, Arizona State University West, United States of America

Introduction

This study describes a metacognitive inquiry process used by teachers as (a) self-reflection to guide their decisions for lesson planning and (b) self-assessment to judge the effectiveness of these decisions for improving classroom practices. Several assumptions guided this study. First, was the importance of giving teachers responsibility concurrent with respecting their abilities and "good will" to improve their own classroom practices. If teachers are given the opportunity to be independent thinkers instead of followers of text-book procedures, the teachers' commitment to teaching will increase.

A second assumption was that teachers tailor their decisions based on their own experiences. Third, teachers are stimulated by awareness of the role they can play in building on their strengths and overcoming their limitations.

There were two major objectives on this study:

- to describe how teachers were trained to use this metacognitive inquiry process to guide their decisions for planning classroom practices;
- to report teachers' perceptions of the use of this approach in the classroom.

More specifically, the following questions were raised: What will be the effects of using this metacognitive approach on the quality of teaching decisions made in the classroom? What will teachers learn about themselves as decision makers? What will teachers learn about students? What changes will teachers make in the classrooms?

Method

Subjects

Participants were 20 elementary school teachers (teaching fourth to sixth grade levels) from two different school districts located in a metropolitan are of southwestern the United States. This group of teachers was selected based on classroom observations and interview results. The results of these classroom observations and interviews indicated that the teachers exhibited four common characteristics:

- These teachers were working in isolation most of the time.
- They did not show innovation in the classroom (e.g., they taught by using a text book almost step-by-step).

- They provided students a great deal of busy work (e.g., filling in blanks and copying from the book).
- Teachers' classes were composed of students with mixed ethnic backgrounds, varied socio-economic levels, and a wide range of cognitive abilities. All teachers volunteered for the study.

Procedure

The procedure and rationale for training the teachers was derived from metacognitive instruction (Peterson & Comeaux, 1989; Pintrich, 1990, 1988a). Teachers were instructed by the experimenter using role playing and modeling in the use of a metacognitive inquiry process. The following questions developed by the experimenter guided teachers as they planed classroom instruction. Teachers used these questions as they planned classes on a weekly basis collaboratively with their peers of the same grade level.

- Which aspects of this topic are more relevant for teaching (e.g., the civil war)?
- Which teaching skills should I use to organize the content in a clear and meaningful way for all students (e.g., graphic organizers)?
- How will I organize each lesson to involve all students in the learning process (e.g., group discussion, peer teaching)?
- Who needs special attention? Which type of feedback should I provide to help students correct their own mistakes (e.g., constructive criticism by considering the error as a source of learning)?
- How do I create a classroom environment that fosters opportunities for all students to be successful (e.g., providing feedback tailored to individual needs in a motivational way)?
- What will be the indicators of my successful planning that would make students better performers (e.g., how can I predict what will happen in class)?
- What did I learn from last week's planning? How can I correct my limitations from last week (e.g., lack of attention to individual differences)?
- How should I take into account students' classroom participation to assess students' performance (e.g., provide written feedback, give points)?
- What will be the best procedure to assess students' performance on this particular topic (e.g., an essay, a display on the cardboard)?
- What did I learn about my performance and the performance of my students from last week (e.g., moving around the room when students were engaged in seatwork, helped students to better concentrate on the task)?

Table 1

Knowledge' Gained by teachers about themselves as decision Makers, Their students as learners, and changes made in classroom practices.

Frequent themes cited	Verbatin excerpts from teachers' journals
Changing of teaching from content-covered to student-centre	"Before I used these questions to ask myself how to do my lesson plans I used to follow the teacher's manual for the students' text-book. It was easier and faster to cover content that was what I most cared about. Now I question myself about the effectiveness of everything I do, e.g., the relevance of what the book says about the content I have to cover, and I found myself very often complementing my instruction looking for other sources." "...When planning the class I question myself about the relevance of the topic. For example, how much students know about the topic, how can I motivate them to learn it? What is the value of teaching it?"
Frequent themes cited	**Verbatin excerpts from teachers' journals**
Team-building as a trustful peer support environment	"In planning classes with my peers we were there to support each other and explore new ideas without being judgemental." "...I notice the value of new ideas getting implemented and applauded by my peers rather than set aside and left untested."
More awareness of students' creativity	"I used to load students with individual seatwork (e.g., filling the blanks). Now I use more active participation and more cooperative work and less direction on how to do it. Very often students amazed me with their creativity in doing displays. For example, the one they did on arranging a zoo. I was impressed with how much knowledge they gained in grouping animals according to geographical regions."
Monitor activities to make students successful	"To those students that never participated in class discussions I began to involve them by asking questions that can be answered very briefly and very easily. I engage students in activities that foster students' success." "Students expressed that reading the book I have selected for them was boring. I took some students with me to the library and asked them to join me in making a new book selection."

Table 1, continued

Changes on teaching style. automaticity ownership, self-confidence	"While students were engaged in seatwork I used to be at my desk correcting their homework; now I move around the class helping students focus their attention on what they are doing and giving them feedback when they needed." "When students fail to learn with one method, I try another one." "After a while I did not need to look at the questions, they already are part of my style not only for classroom planning but also as my teaching style. I find my self questioning more often about the use of everything I do for children."
Self-regulated activities by checking for students' understanding	"Very often, when I am in the middle of teaching something, I stop and ask three or four students, some that appear to pay attention and some that I am in doubt of their attention , to explain in their own words the point I am making. If I am not satisfied with their explanation I explain again, mainly when the topic is important (e.g., division with decimals)."

Data were analyzed by identifying the themes most frequently cited on teachers' knowledge about themselves as decision makers, their students as learners, and changes made in classroom activities.

As indicated in in Table 1, teachers enriched their knowledge base by first changing the focus of their teaching. Instead of focusing on amount of content covered as the centre of their teaching, students' needs gradually became their central focus for planning instruction. Second, planning with peers helped team building which in turn facilitated the selection of relevant curriculum and the sharing of resources to fit students' needs.

The training provided a model of collaborative, professional development which may be a powerful challenge to any who claim that standards can be raised simply by increasing central control over the curriculum and by imposing external attainment targets on students and teachers. Consequently, teachers should be given more freedom to reflect on what decisions to make as well as more time for planning classroom instruction collaboratively.

References

Baker, L., & Brown, A. L. (1984). Metacognitive skills and reading. In P. D. Person (Ed.), Handbook of reading research (pp. 353-394). New York: Longman.

Palinscar, A. S., & Brown, A. L., (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. Cognition and Instruction, 1, 117-175.

Peterson, P. (1988). Teachers' and students' cognitional knowledge for classroom teaching and learning. Educational Researcher, 17, 5-14.

Peterson, P., & Comeaux, M. (1989). Assessing the teacher as a reflective professional: New perspectives on teacher evaluation. In A. E. Woolfolk (Ed.), Research perspectives on the graduate preparation of teachers (pp. 132-152). Englewood Cliffs, NJ: Prentice-Hall.

Pintrich, P. R. (1990). Implications of psychological research on student learning and college teaching for teacher education. In R. Houston & M. H. J. Sikula (Ed.), Handbook on research on teacher education. (pp. 814-896). New York: Macmillan Publishing Company.

Pintrich, P. R. (1988a). A process-oriented view of student motivation and cognition. In J. Stack & L. Mets (Eds.), Improving teaching and learning through research. New Directions for Institutional Research, 57, pp. 65-79. San Francisco: Jossey-Bass.

MOVING TOWARDS TEACHING: A MODEL OF TEACHER DEVELOPMENT

Robert G. Elliott, School of Curriculum and Professional Studies, Queensland University of Technology, Brisbane, Australia

Introduction to Problem

The image of movement is to be found in various conceptions of teacher development. For example, Shulman (1987) has proposed a cyclic process of pedagogical reasoning and action involving comprehension, transformation, instruction, evaluation, reflection and the formation of new comprehensions. Using such a process of pedagogical reasoning, movement in teacher development may be conceived as a cyclic process involving thought and action.

In a like manner, Munby (1986) introduces the image of movement when he interprets teacher action and development in terms of "lessons as moving objects"

The aim of the paper is to construct a model of teacher development in terms of the reported thinking processes adopted by student teachers during their training year. Based on the data collected from student teachers it is anticipated that the model should illuminate the image of movement in teacher thinking and development.

Method

The study is grounded in the narrative mode and such a mode contrasts with the paradigmatic mode of thought (Bruner, 1986) which underpins more conventional analyses of learning and development. The methodology adopted is similar to that employed by Clandinnin and Connelly (1991), Gudmundsdottir (1991) and Egan (1986) in the analysis of narratives.

Eight student teachers were asked to report on the processes by which they developed their professional knowledge during their preservice course and their narratives are taken as the raw data of the study. The narratives were analysed by the researcher to identify dominant characteristics of development.

Results

Three types of elements, regarded as important in describing the student teachers' development, were found in the student teachers' narratives. They were metaphors of development; contexts of development and personal predispositions.

Of greatest significance were the metaphors of absorption, collation and filtering described by the student teachers. These metaphors describe, in the language of

the student teachers, a series of thinking processes employed by them as they attempted to learn about teaching.
In a similar manner the students describe a series of contexts in which learning occurs. In particular, they note contexts of being alone, formal learning groups, informal learning groups and practice teaching as the most significant contexts where their thinking processes are operationalised.
Personal predispositions such as willingness to change, fitting into the system and knowledge from specific contexts were also seen as important in that they influenced the extent to which particular thinking processes were evidenced.

Conclusion and discussion

These data are interpreted in terms of Jacobson's two dimensions of speech analysis- the vertical dimension representing constituent elements of speech and the horizontal dimension involving the combination of these constituent elements to create extended forms. Following Gudmundsdottir (1991), a model for understanding teacher thought involving a vertical dimension (consisting of individual knowledge elements about teaching) and a horizontal dimension (consisting of combinations of these elements) is proposed.
Movement in teacher development and thinking is thus conceived in terms of progressively absorbing individual ideas about teaching (becoming part of the vertical dimension) and later collating them (through a horizontal dimension) to form new thoughts and plans for action. More importantly with regard to movement in teacher development, these thoughts and plans are subsequently filtered in the light of action to produce new ideas which are absorbed on the vertical dimension. This model can be compared with Shulman's (1987) the notion of progressing cycles in teacher thinking. It highlights the constructionist aspect of teacher development and is one which captures the image of movement in terms of the student teachers' reported metaphors of learning about teaching. Further, such a model addresses the need to incorporate the idea of "situated cognition" in theories of teacher thinking. Greeno (1989a: 135) argues that recent research suggests that "thinking is an interaction between an individual and the physical and social situation" and so we should regard knowledge as the potential for situated activity. Such a view of knowledge is compatible with the idea of movement in images of teaching and teacher development because, as Greeno (1989b: 286) notes such a revitalised concept of knowledge would be analogous to the concept of motion in physics.
The model is useful because it is developed in terms of the language of the participants and may be used to encourage students to reflect on their own development and enable them to articulate a personal theory of development. In this sense, the processes of "storying and restorying" suggested by Clandinnin and Connelly (1991) may be able to be set within a theoretical framework of teacher development.

References

Bruner, J.(1986). Actual minds, possible worlds. Cambridge: Harvard University Press.

Clandinnin, D. J., & Connelly, F. M.(1991). Teacher education as narrative inquiry: Narrative inquiry as teacher education. Paper delivered at the Annual Conference. Melbourne: Australian Teacher Education.

Egan, K. (1986). Teaching as story telling. Chicago: University of Chicago Press.

Greeno, (1989a). A perspective on thinking. American Psychologist, 44(2), 134-141.

Greeno, J. G. (1989b). Situations, mental models, and generative knowledge. (Ch. 11) In D. Klar & K. Kotovsky (Eds.), Complex information processing: The impact of Herbert A. Simon. Hillsdale: Lawrence Erlbaum.

Gudmundsdottir, S. (1991). Story-maker, story-teller: narrative structures in curriculum. Journal of Curriculum Studies, 23(3), 207-218.

Munby, H. (1986). Metaphor in thinking of teachers: An exploratory study. Journal of Curriculum Studies, 18(2), 197-209.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57(1), 1-22.

TEACHERS' EVALUATION OF TRAINING AND IMPLEMENTATION OF A NEW FIELD OF STUDY IN THE CURRICULUM

Cees Bruinsma, & Piet van der Jagt, Vakgroep Pedagogiek, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Introduction

In 1985 the State Secretary of Education stimulated the national institutions (CPS, KPC, APS, SLO and CITO) to start projects to contribute to the bottle-necks in primary education with respect to the form and content. Explicitly the development and testing of educational programs for some fields which are important for the realisation of the 'Law of Primary Education' (Wet op het Basisonderwijs, 1985) can be developed (Advies Onderwijsraad, 1986; Lieshout, 1989). In cooperation with CPS and a regional education and school advisory service (GCO) the development and implementation of educational programs for 'religious and cultural education' (Geestelijke Stromingen) on three protestant primary schools was started in 1986 (Projectplan, 1986).

The project was evaluated by researchers who also participated in a group which coached the project till 1990 (Bruinsma & Van der Jagt, 1990).

On base of the project–outlines the aim of this study is:

- to investigate the teachers' perception of the priority of the several topics in the new field of study;
- to investigate the teachers' perceptions of the priorities the team of teachers and the parents of the schoolchildren will give to the topics which will affects the choices;
- to evaluate the implementation of the new field of study in the curriculum (See Andree, 1987).

Method

In this study three protestant primary schools with twenty–one teachers were interviewed in the beginning of the project and at the end. They also filled in a questionnaire with the following topics:

- a hypothesized content and meaning of the new field of study;
- evaluation of the relationship with the own philosophy of life;
- relationship of the new field of study with the curriculum;
- knowledge and the expectations of the new field of study.

867

In the questionnaire the own perception and the perception of the other teachers in the school and the parents of the children were asked which will affect the choice of the topics in the new field of study. The topics which could be chosen were investigated in the interviews. In the questionnaires the teachers could give their priorities for the several topics. In social science research, the measurement of variables is most often done by the ordinal level, while the analysis techniques employed require measurement at the interval level. In this study a continuous scale for measuring variables was used. The teachers and parents were asked to compare the several topics out of the new field of study with a standard presented on a card in random order. They took a card with one of the topics and compared it to the standard one. They wrote their judgment on the questionnaire form and rated the sample by magnitude estimation and an ordinal judgment (Bruinsma, 1987).
The extent of commitment of the teachers with the implementation of the new field of study in the curriculum was measured by a version of the 'Concerns– Based Adoption Model (Van den Berg & Vandenberghe, 1981). In this model seven stages are distinguished in the development teachers undergo when they change their schoolpractice. The instrument consist of 52 items which are measured on a eight–points scale.

Results

The reliability for the evaluation of the topics of the perception of the teacher (median R_2) is .933 for the perception of the opinion of the team .983 and for the perception of the opinion of the parents it is .954. Also the relationship between the teacher's own opinion and the perception of the opinion of the other teachers is high (R= .85). The relationship between the own opinion and the perception of the opinion of the parents is also rather high (R= .73). The conclusion is that in general the opinion of the teachers is conform the opinion of the other teachers in school. There are some differences between the perception of the opinion of the parents by the teachers and the opinion of the parents, but in general they are conform what teachers expect the parents will answer.
In Figure 1 the magnitude scores of the preferences for the topics of the teachers (1986 and 1990) and the scores of the parents of the schoolchildren (1986) are shown.
Some differences can be seen between the scores of the preferences of the teachers in 1986 and 1990.The teachers prefer the topics which correspond with their own convictions. The same pattern can be seen in the scores of the parents. The implementation of the new field of study in the curriculum was successful. The teachers showed more commitment at the end of the project and were able to teach the new field of study to the children (Van der Jagt, 1991).

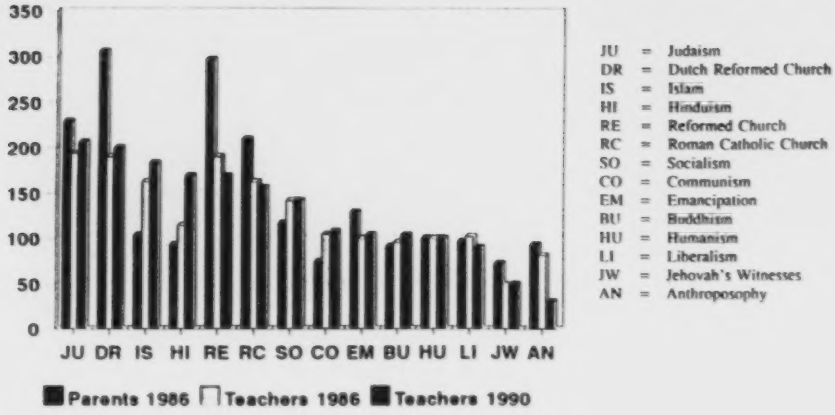


Figure 1
Preferences for the topics in religious and cultural education of teachers (N=13) and parents (N=121).

References

Advies Onderwijsraad. (1986, October). Uitleg. (Projectplan Geestelijke Stromingen CPS : OR 1/224 P.)

Andree, T. G. I. M., & Steegman, P. D. D. (1987). Religieuze socialisatie. Utrecht: RUU.

Berg, R., van den,& Vandenberghe, R. (1981) Onderwijsinnovatie in verschuivend perspectief. Tilburg: Zwijsen.

Bruinsma, C. (1987) Systematic errors in survey–interviews. Amsterdam.

Bruinsma, C., & Jagt, P.J., van der (1990) Evaluatie van de innovatie van een nieuw vakgebied. In J. H. Gerritsen (red.), "Geestelijke Stromingen" voor het protestants–christelijk basisonderwijs: Hand– en ideeënboek. (Hfdst 8), pp. 1-95. Hoevelaken: CPS.

Jagt, P. J., van der (1991). Geestelijke Stromingen in pedagogisch perspectief. Amsterdam,: VU–PPW.

Lieshout, W. C. M., van (red.). (1989). Rijkdom van het onvoltooide. 's - Gravenhage: Ministerie van Onderwijs en Wetenschappen, DOP.

Memorie van Toelichting op het Wetsontwerp Basisonderwijs. en

Memorie van Antwoord op het Wetsontwerp basisonderwijs. (1977). WBO met aantekeningen. Alphen aan de Rijn: Samsom.

Secretariaat Diepteprojekt Geestelijke Stromingen, Christelijk Pedagogisch Studiecentrum. (1986, Mei). Projectplan Diepteproject Geestelijke Stromingen. Eerste projektfase 1986–1988. Hoevelaken: Auteur.

TEACHING AND TEACHER TRAINING

Joost Lowyck, Centre for Instructional Psychology and Technology, University of Leuven, Leuven, Belgium

Introduction

During more than fifty years scholars in the field of education eagerly try to define teaching and teacher education effectiveness. Their endeavours often bear a likeness to the myth of Sisyphus, a king of Corinth condemned forever to roll a stone up a hill in Hades only to have it roll down again on nearing the top. Indeed, although researchers used many theoretical frameworks, models and methods reflecting how teaching and teacher effectiveness should be understood they never reached a univocal nor definite answer.

In this address different paradigms, expectations and outcomes of research on teaching and teacher education are examined trying to grasp its essence and influences.

Research on teaching

In research on teaching, some main phases can be distinguished. However, since an over-emphasising of peculiar characteristics leads towards segmentation, a more integrative perspective is taken.

Effective teaching and teacher's characteristics

During the twenties and thirties, at the times of test psychology, effective teaching is defined as any activity if conducted by 'good' teachers. Effective teaching behaviour in this conception spontaneously emanates from a teacher's personality in terms of stable 'traits'. It is expected that if psychologists can identify the main characteristics of a well-functioning teacher a powerful indicator for effectiveness has been found. Moreover, in times of a sufficient teaching work force, allocation to this stable job is realised by selection, not by training.

The recruitment and education of 'democratic' teachers in the USA was in line with the anti-authoritarian ideology in its reaction to the fascist political scene in Europe around the second World War and, thus, became a paramount object of study. This approach reflects earlier teacher's characteristics studies and anticipates on the future process-product studies in the fifties and sixties.

Process-product studies: teacher's classroom behaviours

The lack of differential outcomes from teacher's personality studies as well as the dominance of behaviourism brought about an orientation towards teaching 'behaviours'. Moreover, the golden sixties with an enormous growth of schooling due to both the post-war baby boom and the democratization power of secondary education necessitated an increased number of teachers. Because selection seemed no to warrant effectiveness emphasis, then, is laid on training.

The basic question is which teaching behaviours in classrooms (process variables) are effective as to bring about learning gains in pupils (product variables). The complex teaching activity is fragmented into discrete teaching skills in line with the analytical paradigm. Some powerful process variables are extracted from meta-analytical reviews. The problem with this approach, however, lies in the analytical and formal nature of the process variables, independent from subject-matter and student's characteristics. As a solution to the analytical deficiency, researchers try to 'chunck' elementary teaching behaviours into more encompassing models, like 'active' or 'adaptive' teaching. Nevertheless the direct relationship between teaching behaviour and pupil achievement remains problematic. Consequently, it is hypothesized that many 'intervening' or 'mediating' variables determine effective learning.

Mediating variables : learning time and pupil cognitions

Since the direct link between instructional input (teacher behaviour) and output (learning outcomes) seemed an insufficient effectiveness paradigm the focus is now on the throughput: the variables mediating between teaching and learning. Studies on 'Aptitude-Treatment-Interaction' as well as the 'Beginning Teacher Evaluation Study' revealed the importance of individual processing differences for active learning. Moreover, differentiation and individualization were regarded as vehicles for the solution of problems typical of mass education.

In a first approach time-on-task is considered to be a powerful predictive variable though researchers gradually experienced difficulty with the concept itself, since it is a mere quantitative measure. In the search for more qualitative mediating variables, two categories are distinguished : the cognitive and the social one. Whereas the former takes socialisation as the main intervening process, the latter concentrates on cognitive processes of pupils.

Teacher thinking

The shift in the psychological scene from behaviourism to cognitivism, along with the refinement of qualitative methodology in the seventies paved the way for the study of teachers' cognitions. Teaching is no more considered as the sum of isolated competences, but as a profession with a high load on cognitive skills due to the complexity of both the educational environment and the teaching task. It is a job similar to the one of physicians, engineers and architects with emphasis on judgment, problem-solving and decision-making.

Although most studies of teacher's thinking are explorative and conceptual in nature, they nevertheless offer some indications as to effective teaching. At the beginning of this line of research attention is paid to the description of the 'naturalistic' cognitive processes of teachers during the pre-active, interactive and post-interactive phase of teaching. Later on, the interaction between phases

became object of study. It is observed that the 'psycho-logical' way in which teachers function often contrasts with 'logical' teaching models advocated by instructional designers and teacher trainers.
Because descriptions of cognitive processes were content-free, the study of teacher's subjective or implicit theories (teacher beliefs, teacher's practical knowledge) focussed on the very content of teacher cognitions. It is shown that teachers do not use a complete nor clear model of teaching, but an eclectic mix of rules of the thumb, generalisations, opinions, values and expectations. An important issue is how the teacher's cognitive understanding and use of subject matter content can be elucidated. Any unnecessary isolation of types of teacher's knowledge is avoided focussing on the intrinsic interaction between subject-matter knowledge, pedagogical knowledge and curricular knowledge.

Research on 'learning to teach'
From an intruding analysis of the methods and content of teacher education some important characteristics come to light.

Theory and practice
One of the paramount issues in teacher education rests on the interpretation of the link between theory and practice. During a long period of time, theory was isolated from practice. The focus was on the theoretical enrichment of future teachers in view of the subject they have to teach, while practice was restricted to the observation and imitation of 'expert'-teachers in real classrooms.
Later on, the dichotomy between theory and practice was severely criticized by the so-called 'competence-based' teacher education. The foundation of any training was sought in an extensive task analysis from where all the necessary knowledge, attitudes and skills were derived and organised in the curriculum. However, the discrepancy between the fixed inventory of tasks and the dynamic evolution of education itself became a major problem.
Because in recent times more emphasis is laid on the flexibility of teachers and on their cognitive processes as well, the basic skill of teaching is perceived to be 'reflected practice'. Teaching is not the mere knowledge of subject-matter complemented by practical tricks, but 'knowledge-in-action' taking the complex situation into account.

Training methods
It is only in the sixties that more systematic, behaviour oriented training methods are instrumented. Examples are micro-teaching, mini-courses, interaction-analysis. These methods are in line with both the dominant process-product paradigm and the competence-based education with an almost exclusive focus on partial teaching behaviours.
When cognitive processes of teaching became more important objects of study, emphasis is laid on simulation, case-studies and analysis of protocol materials. Moreover, the behaviourist methods are reconsidered and transformed into more cognitive tools, like the cognitive approach of micro-teaching.

Discussion

From the analysis of studies on teaching and teacher education, some main observations can be made.

Firstly, within both research on teaching and on teacher education, there is a development from personality characteristics to teaching behaviours and teacher cognitions. These reorientations bear on extrinsic influences, like the current values of the society, the educational scene and the research paradigms.

Secondly, the empirical knowledge-base for decisions about effective teaching and teacher education remains deficient. The effectiveness criteria seem to evoluate from extrinsic measures of the teaching output in terms of learning gains towards intrinsic teacher cognitive processes.

Thirdly, new orientations on the teaching profession are influenced by the state-of-the-art in educational psychology and by the demands on teachers in terms of societal needs. It is clear that teachers need to be very flexible and responsible professionals, due to the rapid changes in the educational goals. Creating powerful learning environments, stimulating higher-order cognitive processes in pupils, enabling transfer of knowledge, elicit cognitive flexibility and foster ethical awareness are considered of paramount importance for the survival of society. New research endeavours are needed in order to empirically grasp the necessary conditions for the new professional qualifications and for their instrumentation as well.

References

Doyle, W. (1986). Classroom organization and management. In M. C. Wittrock (Ed.), Handbook of research on teaching. (3rd ed.). (pp.392-431). New York : Macmillan.

Dunkin, M. J., & Biddle, B. J. (1974). The study of teaching. New York : Holt, Rinehart & Winston.

Gage, N.L. (1978). The scientific Basis of the art of teaching. New York: Teachers College Press.

Rosenshine, B., & Furst, N. (1973). The use of direct observation to study teaching. In F. M. Travers (Ed.), Second handbook of research on teaching, pp. 22-183. Chicago: Rand McNally.

Shulman, L. (1986). Paradigms and research programs in the study of teaching: a contemporary perspective. In M. C. Wittrock (Ed.), Handbook of research on teaching. (3rd ed.) (pp. 3-36). New York: Macmillan.

AN EVALUATION OF A NATIONAL TRAINING PROGRAM ON EDUCATIONAL LEADERSHIP FOR PRIMARY SCHOOLS

C. Matthijsen, H.P. Brandsma, & W.G.R. Stoel, The Netherlands

Introduction

Since the start of the eighties the interest of both researchers and policymakers in the organization of schools and the role of schoolmanagement therein has increased considerably. Research on effective schools had led to the conclusion that the role of the principal in the effectiveness of schools is a crucial one. Especially the shift from a more administrative leadership role towards educational leadership proved to be effective. However the management of primary schools was not always adequately equipped to make this shift. This has resulted in an increasing need to train principals towards an educational type of school leadership. Therefore the central Dutch government has initiated a national 5 day training program for all (some 8000) principals in primary education. With regard to these refresher courses the three Pedagogical Centres (KPC/CPS and APS) have developed the curricula that can be used in the actual training. The training itself is given by the Teacher Training Colleges (PABO) in cooperation with the regional educational support agencies (SBD). The intention of the national policy was that all principals should attend this training within a four year period, starting in 1988.

Educational leadership

Apart from all kinds of administrative tasks within the framework of general schoolmanagement the principals also have an important task in the development and supervision of the so called innovative educational policy at the schoollevel. Within this framework of strategic management it is important that principals come to see it as an integral part of their task to:

- develop long term educational goals;
- initiate,implement and support educational innovations;
- coach their staff by means of intervision and supervision;
- integrate the curriculum between teachers and schoolyears, and to
- evaluate and analyze the progress of pupils.

The principal takes the role of an internal initiator and supports the school development process. The training program should contain all these elements of educational leadership.

The evalution study

As an external part of the national training policy an evaluation, both formative and summative, has been conducted by the Centre for Applied Research in Education (University of Twente). This evaluation is intended to give a description of the form and content of the training program, the satisfaction of the participants and to evaluate possible effects on management behaviour of principals. Within this evaluation study two research questions are central:

- Which are the characteristics (with respect to form and content) of the training program and are these in accordance with the policy goals of the central government?
- Has the participation of principals (the trainees) created the intended shift towards educational management behaviour and is this shift also visible for the teaching staff?

This second research question will be answered in this paper.

With regard to the possible effects in terms of a change in management behaviour both a random sample of principals and three members of the teaching staff received an evaluation form on which they could describe the change in educational management behaviour of themselves (trainee principals) respectively the behaviour of their principal (staff members) after one year of ending the training program. For this evaluation an instrument has been constructed in which all dimensions of educational leadership behaviour are incorporated. These instruments proved to be adequately reliable.

In order to control for possible time effects the same questionnaire has been sent also to a random sample of principals who had not received this training and three of their teaching staff members (n= 250).

By means of this controlled, retrospective pre– and posttestdesign with both principals and their staff as respondents it is possible to trace the (magnitude of the) effects of change in educational leadership behaviour.

Results

The result of the analyses with regard to the first research question show that most T.T.C.'s have not restricted themselves solely to aspects of educational leadership but have also incorporated aspects of general schoolmanagement in their training program. On average they spend about 25% of their time on general management issues. As there had already been a national refresher course in general management one can conclude that the content of these courses don't coincide completely with the policy intentions.

Trainee principals are reasonably satisfied with the form and content of the refresher courses.

With regard to changes in educational management behaviours one can conclude that in general the trainee principals have changed their behaviours more than their non–trained colleagues. This shift in management behaviours is also visible for there staff members, although they estimate the changes to be less than the principals themselves.

As a general conclusion one can state that this national training program has been successful in the sense that principals have changed their management behaviour towards educational management and that in general their teachers agree with them and also experience this shift.

THE IDENTIFICATION OF A COMMON CORE CURRICULUM IN THE TRAINING OF TEACHERS, SOCIAL WORKERS, HEALTH WORKERS AND COMMUNITY EDUCATORS

Michael Hough, Senior Lecturer, Jordanhill College of Education, Glasgow, Scotland, United Kingdom

Increasingly Teachers, Social Workers, Health Workers and Community Workers are being asked to collaborate and work jointly across the grain of their professional traditions. On the ground, in the classroom, clinic, community centre and social work offices, they are finding that they are often dealing with the same issues and the same clients. They are having to develop strategies for professional demarcation and collaboration, but their training has not equipped them for these new realities nor are their management systems tuned to promote collaboration.
Our research looked at five examples of projects where professionals including Health Workers, Teachers, Social Workers and Community Educators were employed to work collaboratively and establish new relationships with each other and with the communities within which they worked. Our methodology was to visit each project and conduct structured discussions with project workers. We then invited them to submit written evidence of their work and to come together to share and explore their experiences. Specific training sessions for some of the projects were also organised, testing out the processes identified as potentially useful in promoting more effective working practices. From this we were able to obtain a wealth of data in various forms about the nature of collaborative practice.
Using this data together with a review of some of the U.K. and American literature, we were able to identify certain focal activities which we regarded as common to a range of professions. Our conclusion was that these core activities provided a potentially fruitful area of common ground that could radically change the nature of pre-service training and contribute to a re-focussing of inservice work.
Element of this core curriculum include:

- Objective setting and task analysis
  All professional groups have moved closer to articulating their purposes and objectives. Although the outcomes of this work may be different, the process of defining objectives and identifying work tasks is a common one and draws on a common body of knowledge.
- Creating learning environments
  Attention to learning and the acquiring of new skills is no longer the preserve of the educationalist; other professional groups now acknowledge that they are also involved in the learning process, although their words to describe such activity

may vary, for instance empowerment, raising consciousness. At the same time teachers are beginning to use methodologies derived from community work and counselling to enhance their repertoire of teaching techniques.

- Assessment and evaluation
  Both words tend to mean different things to different groups, but the underlying trend is towards all professional groups engaging in a process of assessment and evaluation at some point in their professional practice.
- Monitoring and appraisal
  Accountability has been a key word in the 80's. Teachers, Social Workers, Health Professionals have all been, and are currently, engaged in refining and developing appropriate systems of appraisal and monitoring of achievements and performance.
- Staff development
  All professional groups have a commitment to staff development. Sometimes the resources have not been available to make this a potent force for change but what is emerging is that there are common elements of staff development, whatever the professional base.
- Supervision, consultancy and counselling
  Again these are terms that are used in different ways by groups, but to a greater or lesser extent the practice that underlies the terminology is common to all professions.
- Relationship to client groups
  Whether we regard our client groups as customers, shareholders, partners, or just local people, the issue of clarifying this relationship and understanding the forces at work in such relationships, is central to each professional group.
- Community involvement
  Whether through involvement as voluntary members of management groups or through more complex relationships, each professional group acknowledges that it does not operate in a vacuum and therefore understanding and relating to the community within which they work is a crucial factor in determining the outcome of practice.

The other outcome of our research was to identify and, in most instances, pilot a range of training procedures designed to promote greater and more effective interprofessional collaboration.

Our experience suggests:

- that people welcome the opportunity to meet and share information if it is presented as non-threatening and if it allows people to share information from various aspects of their lives, not simply the professional arena;
- that a training programme needs to be negotiated with those taking part so that they have a sense of ownership of its content and process;
- that people prefer to work on material that is derived from the reality of their practice;
- that moving towards the creation of action plans to resolve some of the day-to-day practicalities is beneficial;
- that the sequencing of training activities requires careful preparation;
- that time for evaluation of the impact of training structures needs to be built into the process from the beginning;

- that staging the training over a period rather than a one-off event is more productive and likely to lead to more significant changes in practice.

One example of a training activity is a *prioritisation exercise*. This involves a number of key stages from which participants identify a range of options in relation to some problem or task. These are then reproduced on cards and participants invited to arrange the cards in a diamond formation reflecting their relative importance. Further work is then done on the basis of this, in small groups, looking at what is within the participants' locus of control. Finally, action plans are derived so that there is a tangible outcome to the exercise.

Another of the cornerstones of our work was the creation of case studies derived from the experience of the project workers. Our approach was to listen to the workers' accounts of their experiences, then to write these up and give them back to the projects for clarification. Finally they would be altered to ensure confidentiality and thus allow them to be used by other groups to gain insight into key practice issues. These case studies represented critical incidents in professional practice. When used in a training programme they have proved valuable stimuli to focussed and reality-based discussion of important and complex issues of collaborative work.

Through the identification of elements of a core curriculum and the production of a variety of related training materials, this research has made a small but significant contribution to the discussion and practice of interprofessional collaboration in the U.K. The opportunity of extending this to a European context opens up new and exciting possibilities in a field that will be particularly relevant in the cost conscious and integrated world of the 1990's.

Copies of the full research report entitled *Working Together* by Mike Hough and John MacBeath (Jordanhill College 1991) can be obtained, price £17.00 (including postage and packing) from Mike Hough at Jordanhill College, Glasgow, Scotland.

878

ACCESS & TRANSITION INTO TEACHER EDUCATION IN AN INCREASINGLY ETHNIC DIVERSE SOCIETY

John R. Meyer, Faculty of Education, University of Windsor, Windsor, Ontario, Canada

The presentation will report on a Canadian provincially funded research project that (a) investigated the current status quo of representation by minority candidates in Ontario teacher education, (b) used survey and interview methods to determine the needs and reactions to current access and transition programs for said candidates in three tertiary institutions, (c) by means of a series of recommendations, proposed changes in the admission and transition policies and/or procedures. Upon completion of the analysis of the data, attention was devoted to initiating changes through the sponsorship of an invitational conference to discuss the findings and action proposals. There is some optimism that suggested changes will occur particularly with the impending Ontario policy formulation on equity issues.

Formal education, schooling, is an exclusive provincial jurisdiction in Canada. Such division of powers dates from Canada's first constitution, The British North American Act (1867), when religious and educational minority rights were protected for New France and British Canada. Little has changed since that date nor has there been much discussion of that issue in the constitutional reform debates. As Canada continues to shape its future, the movement of populations is having dramatic impact on the pluralistic composition of its population. Such metropolitan centres as Montreal, Toronto, and Vancouver have experienced an influx of immigrants and the growth of descendants of first generation immigrants of neither English nor French origin. Furthermore, it is predicted that by the year 2000, approximately 80% of the new entrants into Ontario will be from one of the four minoritygroups, i.e., aboriginal communities, disabled people, racial minorities, and women. The Province of Ontario, Canada's most populated (9 million) Province, with a capital city of Toronto (3 million), has representation of some 70 ethnic groups. To recognize this diversity and attend to specific cultural and linguistic needs, the governments of Ontario have been active in promoting equity initiatives (1988) and the formulation of a comprehensive equity policy with compliance (1992). The school populations have also been affected by the presence of an increasing number of children of immigrant backgrounds. To respond to this change, a policy on Heritage Languages was adopted and financially supported by the Ontario government. But little has been done to insure that the teaching: corps also reflect the cultural diversity of the students. If it can be correctly assumed that in order to be most effective, teachers must also be from

those ethnic backgrounds of what has become the dominant student population of Oriental, Asian, Black, and East European students in certain Ontario classrooms. But a quick tally by visible observation of our 10 faculties of education over the past decade indicated little more than a 2% presence of candidates for teacher certification. The faculties of education graduate and certify some 3,500 students each year and receive some 25 - 30,000 applications each year. The objectives of the research project were to provide:

- an analysis of the results of data collected from surveys and interviews of a sample population of students enroled in teacher preparation programs at three faculties of education;
- a description of existing models in Canada and the U.S.A. that offer positive promise for improvement in recruitment and retention of under-represented teacher education candidates;
- a model design for transition from secondary and/or tertiary degree programs into concurrent or consecutive teacher preparation programs;
- a series of proactive recommendations directed to the appropriate stakeholders in public institutions and agencies.

Discussion

The analysis of the data indicated that only some 1-2% of the students surveyed were from under-represented groups, i.e., ethnic minorities and native persons. Such a percentage can be generalized to other provincial faculties of education with exceptions for Lakehead, Nipissing, and Queen's at which special programs for native persons exist and for Toronto at which a significant increase of minority students has occurred in 1991-92. One major obstacle in collecting accurate data was the fact that the official application form administered through a central application agency does not provide an explicit opportunity to self-identify on the basis of ethnic or minority background. Only two models (Winnipeg and Toronto) for recruitment and transition were discovered in Canada. Several state and university models were identified from the U.S.A. In that nation, the primary under-represented groups are the Afro-American and the Hispanic or Latinos. Noteworthy system models exist in the states of South Carolina, Oklahoma, and New Mexico. An empowerment, self-actualization transition model from British origin was outlined and recommended at the tertiary level. The seven recommendations address the issues of: special career counselling for targeted groups at the pre-tertiary level revised teacher recruitment policies; financial assistance; prerequisite transition programs for targeted groups; revised application forms; intergovernmental research unit; and a clearing-house or repository for all relevant information concerning ethnic diversity as it relates to the teacher and student representation.

One issue that emerged during the research and during the consultation held for decision makers was the need for much greater match of teacher and student population on the basis of gender. Currently male candidates are sought for the elementary grades. Female candidates are sought for the secondary maths and sciences areas. The status and rationalization of teacher education programs will depend on restructuring. The future of education in Ontario as elsewhere is inextricably tied to the preparation of teachers. Some attention will be given to the restructuring of the entire educational system in the Province of Ontario.

SCHOOLLEADERS IN TRAINING: AN INTERNATIONAL COMPARISON

Q.H.J.M. van Ojen, & S. Karsten, Centre for Educational Research and Departement of Educational Sciences of the University of Amsterdam, Amsterdam, The Netherlands

Introduction

The increase in institutional autonomy, educational innovations and new organizational insights, caused fundamental changes in the tasks and powers of headteachers. One of the mean consequences of this changing role is the increasing need for professionalization in educational management. Therefore, but also in addition of the set of measures issued by the European Community concerning the Europe of 1992, a comparative study aimed at obtaining a picture of the way in which professionalization activities have been given shape in various countries at national as well as institutional levels, is useful. The second aim of this study is to find out whether a certain form of European cooperation in the area of educational management is desirable.

On the basis of a study of the professional literature two different questionnaires were developed and sent not only to central, regional and local authorities, but also to experts and training institutes in fifteen different European countries.

The results of the research are mostly general in nature. This means that it is possible that individual training institutes deviate from national and thus general conclusions. Furthermore, the research was finished at the end of 1990. It is possible that new developments within one country has started since.

Results

The degrees to which the countries investigated aim for a professionalization of educational management, differ widely. They vary from total absence to top priority. But not only does the priority given to the professionalization of educational management differ for each individual country. The ways in which systems of training are organized, particularly by the relevant authorities and other organizations - as well as the division of responsibility - are also different for each country. For example, in Luxemburg the professionalization of educational management is almost completely in the hands of the central government. In Germany and Switzerland most of the responsibility lies with the federal states and the cantons. Finally, in England the emphasis is much more on local authorities,

881

whereas in Sweden all levels of authority have responsibilities and initiate activities.
The overview below, contents a number of models which represent the way in which training systems are organized. All the various countries are included in these models.

centralized model	decentralized model	mixed model
France	Switzerland (federal level)	Sweden
Luxemburg	W.-Germany (federal level)	Norway
Spain	Belgium (local/regional level)	Finland
Italy	England (local/regional level)	Denmark
Northern Ireland	Scotland (local/regional level)	
	The Netherlands (institutional level)	

In the centralized model the emphasis is on central government. It is the government that issues guide-lines, it finances and organizes, it initiates and coordinates. The central government has the responsibility for the development of professionalization activities and keeps a high profile in this matter.
In the decentralized model the roles that are just mentioned are delegated to lower levels of authority. These include regional levels (such as cantons, provincial authorities and the like), local levels and the level of the training institutes themselves.
The mixed model has characteristics of both the preceding models. The focus in the division of responsibility is not clearly on one level, but on several levels. This leads to a complicated picture of the training system as a whole.
It is clear that there are great differences in the professionalization of educational managers, not only in terms of actual commitment but also in terms of organizational and legal frameworks.
However, these differences tend to fade away if we look at the level of institutions. The system and content of educational management training courses as organized in the various countries, are much less divergent. The courses are aimed at teaching leadership, at the organizational and administrative aspects of leadership, as well as at personnel management and other relevant aspects. In almost all the countries it is assumed that these educational management training courses will lead to an increase in school effectiveness. Most of the institutes, regardless of nationality, offer initial and in-service courses for experienced as well as less experienced head teachers.
However, the selection criteria differ for each training institute. There is also a great difference in the number and types of courses offered. This is particularly true at the institutional level and is to a lesser degree dependent on the country itself. As far as certification is concerned, we see a far from uniform pattern. Despite the fact that there are many institutes who issue certificates, they are all non-formalized. This means that international comparison is impossible.
The way in which the training courses are given, consist for most training institutes of a combination of teaching methods, in which theoretical and practical aspects are linked together as much as possible. In order to achieve this linkage the methods that are most often used, come from adult education. The methodological

approach is very varied but above all aimed at practical application. A combination of lectures and small working groups or group assignment can be found in practically all the institutes, regardless of nationality. Individual assignments and discussions are also frequently used as working methods by most institutes.

Conclusion

The main conclusion must be that the professionalization of educational management is still in its infancy in many countries. However, in most European countries the interest in professionalization increased in the second half of the nineteen-eighties. Especially in Northern and Central Europe activities in this area have been institutionalized to a greater or lesser degree.

Although the number of internationally oriented training activities is increasing day by day in the world of industry and commerce, there are few signs of this happening in the area of educational management. Educational managers are not yet ready for a more internationally oriented society. We would therefore suggest that we should lose no time in answering the question what kind of contribution training courses should make to the internationalization of Europe. In spite of the increased interest in professionalization, activities in this area have hardly been brought to bear on the internationalization of Europe.

Discussion

The differences in educational systems, training systems, educational ethos and the professional qualifications required for head teachers, are too great to allow us to enter into international agreements or cooperation at all. We think that the differences are, in fact, too great. Therefore, it would be difficult to agree a common core. It would be better to take one or more different routes like bilateral or multilateral exchanges of information and a general effort to increase the transparency of each other's education systems. Differences should not only be allowed for, in some cases they should be positively encouraged. In other words, uniformity of training systems should be discouraged. Not only the differences between cultures but also the differences between the various education systems are simply too great.

Barring uniformity, what other roads are open to us? We can think of two in particular, namely the creation of data-banks and exchange programs. Data-banks with surveys of key qualifications for professions and trades etc. would seem a distinct possibility and particularly appropriate in the context of the professionalization of educational managements. It would also seem important for all institutes concerned with the professionalization of school managers to be able to regularly receive up-to-date information about common problems in the area of educational management and about European education systems. Finally, international mobility on the labour market for head teachers could be increased through the creation of exchange programs. The role of the European Community as the potential coordinator of such programs is certainly worth discussing.

COMPARING MODES OF TEACHER EDUCATION: INSIGHTS INTO THE PROCESS OF A RESEARCH PROJECT

Lan Barton, University of Sheffield; Elizabeth Barrett, & Geoff Whitty, University of London; John Furlong, University College of Swansea; Sheila Miles, Hammerton College, Cambridge, United Kingdom

The past few years have seen an unprecedented number of attempts to find new ways of training teachers, both by reforming existing routes and by developing new ones. Accompanying these changes in England and Wales, there has been a vigorous public debate concerning the value and relevance of teacher education.

The authors of this paper have received funding from the Economic and Social Research Council to compare the nature and costs of the various modes of teacher education that currently exist in England and Wales. These include some courses that are based largely in institutions of higher education and other approaches that are primarily school-based.

The research being carried out by the authors aims to:

- provide a sharper focus to current policy debates about the nature and costs of initial teacher education;
- provide an up-to-date database on the nature of a range of approaches to initial teacher education and to identify the competencies and models of professionalism that they seek to engender;
- analyse the direct and indirect costs and benefits of the various modes of teacher education;
- identify the dimensions that need to be considered in future research on the nature, experience and outcomes of teacher education.

The research team has conducted a postal survey of all courses of teacher education in England and Wales and is currently engaged in a more detailed follow-up of a representative sample of 50 courses drawn from across the various modes. Interviews are also being conducted with the major stakeholders in teacher education.

An Interim Report (1992) of the first phase of the project has now been produced and circulated to the institutions involved and a Conference to discuss the findings has taken place.

The paper will consider some aspects of the research process including, the motivations of the researchers, becoming a research team, making decisions over the design of the project and creating a data-base. Specific examples of dilemmas and outcomes will be highlighted. Finally, the paper will reflect upon some of the

dilemmas that arise from carrying out research of this type in a context where the future of teacher education has become a highly-charged political issue.
This paper is thus a personal account of the experience of undertaking a collaborative research project. It takes seriously the importance of reflection on the research process in which difficulties and achievements are acknowledged and discussed.

References

Barrett, E., Barton, L., Furlong, J., Galvin, C., Miles, S., & Whitty, G. (1992). Initial Teacher Education in England and Wales: A Topography. Copies are available for £1.50 from the Faculty of Education, Goldsmiths' College, New Cross, London, SE14 6NW.

885

THE 21ST CENTURY ROMANIAN TEACHER - A PERSPECTIVE VIEWPOINT ON HIS PROFESSIONAL DEVELOPMENT AND MORAL CHANGES

Anca Dumitrescu, Institute for Educational Sciences, Bucharest, Romania

In a society undergoing so many outstanding political, economic and social changes, it is rather difficult to foresee the future evolution of the education system in general, as well as that of its basic components, in particular.
In today's Romania, the Government and the whole people are aware of the fact that education represents one of the main national priorities. Its general reform against the existent background of deep-going changes involving all the country's activity fields, will bring about important mutations in the content and organization of the Romanian school-system.
How are the teacher's status and behaviour modifying as a result of the new outlooks, demands, concerns and tasks? To what extent will his professional development and moral profile be influenced both by the general changing background of nowadays Romania and by the instructional framework with its peculiar directions of evolution?
We should think of the fact that the teachers who are now taking part in the design and implementation of a new education system structured on democratic bases will work together with other younger generations of teachers in the Romanian schools at the beginning of the 21st century.
Our study will analyze the professional and behaviour model of tomorrow's school teacher trying to anticipate and describe the configuration of his evolution paradigm. Thus, we shall discuss important problems, such as: the teacher's greater concern for developing his general scientifical and technological competence, alongside with his knowledge of the specific school-subject he is dealing with, his democratic views on instruction content and organization, an increasing receptiveness to the open dialogue with pupils, parents and school leading staff, direct interest and involvement in school life, his growing awareness of Education's European dimension and practical efforts for achieving it, a.s.o.
All these main directions of the teacher's future development represent important aspects included in the dynamics of teachers' basic and further training in Romania, which should be deeply restructured according to the demand of nowadays democratic society.

The important mutations that have occurred in the Romanian teacher's moral and professional profile after the 1989 Revolution and will continue to develop at the beginning of the next century directly influence the organization and content of teachers' training in our country.

In this respect, we shall point out some of its main priorities: to establish a better relation between theory and practice, between the contents corresponding to various scientifical, psycho-pedagogical, methodological competences; increasing the practical-applicative character of training activities, the predominance of topics linked to psycho-pedagogy and school-subjects methodology; assuring coherence and unity within the diversity of activities for teachers' training-common minimal programs and alternative modules of a great variety, peculiar to educational situations and the real need for professional development; a more efficient correlation of the training programs content with the requirements of the new curricula for primary, secondary and technical-vocational education; the initiation of the teaching staff in forming and developing attitudes and aptitudes (character traits, receptivity, creativity, flexibility, etc.), a more efficient style of intellectual activity.

We are all aware of the fact that the success of the General Education Reform and the creation of a new democratic school-system in Romania will greatly depend on the competence and responsibility of today's and tomorrow's teachers.

INITIAL FORMATION OF TEACHERS AT SECONDARY EDUCATION LEVEL IN SPAIN: AN UNSETTLED CHALLENGE

Vicente Fernandez Benítez, & Johanna Mª Poulussen van Balen, University of Cantabria, Spain

Secondary education is in a permanently changing situation
Not only in Spain, but also in almost every Western–European country. Also, attending to this problem the professional qualification of the teacher in secondary education level doesn't seem to receive the same attention from the different ministries of Education, which debate between two extremities:

A consideration of the teacher as a "craftsman" (teaching is the way to learn how to teach)
The theory is something which can be made up for with the practice in the classroom. The reform in the initial formation of teachers in secondary education level which is pretended to be introduced in the United Kingdom nowadays seems to get very close to this view. I want them (Trainee Teachers) to concentrate on classroom skills rather than academic training... I'm interested in real children, in real classrooms, not academic theory (Kenneth Clarke, the Education Secretary of U.K., in The Observer, 5th of January of 1992).
During the 1970's and 80's in Spain a formation model was developed based on the learning of teaching skills. The dominion of certain teaching strategies was considered basical for effective teaching (Villar Angulo, L.M., 1980; 1982). According to these ideas formation methods were derived that were quite specific, and some of them very constructive, like microteaching.

An "academical" conception of the initial formation of the future teacher in secondary education level
During his period of formation, it's the theory which takes the main part of his time, although there exists a period of practice at schools. On the contrary of the former empiricism, we find ourselves with a sum–up of theoretical courses that are rather distant from practice. It's perfectly comprehensible that there are some risks: the extent of the curriculum and the little time dedicated make it impossible to be have a study with serenity and sufficient depth. They pretend to teach in one or two courses what a pedagogue took to learn during all his university career. So there are some speed–courses on almost every psychopedagogical subject; the classes estimulate a passive learning, while everybody agrees about the need of an active learning; the "standard" problems the students find on their way result steady, with

little possibility of new ways of reasoning facing new situations. The practice at schools, under the direction of a tutor are nearly always some kind of an imaginary testfield, where it's hardly possible to apply any theory learnt at university nor is it possible to be responsible of the action during the short period of practice. This is the model, which, with some differences in each country, dominates in the main part of the Western–European countries.

Teaching is not only an art, or the result of experience, but also, apart from that, a profession which needs a previous knowledge of a specific theory and an initial practice

That's why the initial formation of the teacher at secondary education level is a basic element in whatever reform of secondary education planned, like the one now in motion in Spain. First we will have to know what kind of teachers we want to form in order to create the curriculum we will offer them.

What criterions must the initial formation of secondary level–teachers be based on?

We think the following two judgements should be considered:

- Confront theory with practice. His formation should give him the conditions of being able to operate on the practical field, on specific and real scholar contexts. This means the discarding of an only theoretical formation and, on the other hand, a practical formation based on the establishment of routines with little theoretical basis. Only in this way a teacher will be able to teach a feeling of investigation, making diagnosis of the situations and making the most correct decisions possible (Garcia–Varcercel, A., y Fernandez Benítez, V., 1992; Stenhouse, L., 1984; Perez Gomez, A., 1987).
- Make it possible that the teacher changes his spontaneous thinking about his own activity. This spontaneous thinking is normally occasioned during his large experience as a student (Gene, A., y Gil, D., 1987). Only overcoming this conceptual obstacle, the future teacher will be able to obtain an efficient formation.

A deep knowledge of the subjects to teach

This permits us to free ourselves from being only transmitters of the contents of the conventional curricular materials (Bromme, R., 1988). To make decisions and choose contents that respect the educational necessities of the students, we need to have a deep knowledge of the methodology, the latest advances, its affinities with other material in order to obtain an interdisciplinair teaching, etc. (Furio, C., Gil, D., Pessoa, A., y Salcedo, L., 1992) This deep knowledge is essential to get good teachers en practice.

Some basical psychopedagogical understanding

Referring to the following areas knowledge: evolutionary psychology, learning psychology, theory and sociology of education, general didactics. Also material about the resorts and other organizational help to learning, (groupwork–techniques, new technologies, etc.). Finally, administrative foundation and scholar organization within a frame with European and world–wide references (comprehensive and non–comprehensive education, types of educational administrations, etc.).

Special didactics

All the former knowledge must incise on the scientific basical formation of the future teacher. The central role of special didactics is derived from this idea: every teaching strategy is determined mostly by its contents. departing from the work about the specific materials the teachers will know how to identify the essential aspects and also adopt the educational decisions that have been presented to them in abstract terms (Mc Dermott, L., 1990).

General indications for a change

In Spain the initial formation of teachers is boarded once the students have obtained their university degree. It is a course given by the Institutes of Sciences of Education (I.C.E.) of the respective universities. These courses are called C.A.P. (courses to obtain the Certificate of Educational Aptitude). Parting from the evaluations and the experiences that have been carried out two lines for a reform have arisen:

The creation of specific studies for teachers of secondary level

It's known as the proposal of group XV of university experts, and it has been turned down by the Ministry of Education. It is mainly criticized because these studies are no guarantee for good teachers, who may have little basical knowledge on their material, as occurs at the primary level. Besides, having only one way of being able to get access to this profession could make it impossible for valuable candidates to obtain it (Furio, C., Gil, D., Pessoa, A., y Salcedo, L., 1992).

Studies based upon the reform of the current system

A good theoretical–practical formation could be achieved and a changing in the spontaneous thinking of the future teacher could be promoted if the negative aspects of the C.A.P. were corrected. Besides, this framework respects the Spanish legislative frame (L.O.G.S.E.) and it takes into account other experiences (Krüger, M., y Veugelers, W., 1988). We propose a strategy that includes three operations during the formation:

- During the second half of the university career, theoretical subjects must be included about "Education" and "Specific Didactics".
- One course after having obtained the degree. This course should have psychopedagogical contents, school organization and didactic resorts. The number of students should have to be limited. The main activity would be practice at a secondary level school and seminars at university.
- Assistance to the new teacher during his first year working as a teacher at secondary education level. A follow–up course of help to the beginning teacher (professor novel) which would absorb the first shock with his practical teaching in the classroom (Imbernon, F., 1989).

860

References

Bromme, R. (1988). Conocimientos profesionales de los profesores. Enseñanza de las Ciencias, 6 ,(1), págs. 19–29.

Furio, C., Gil, D., Pessoa, A., & Salcedo, L. (1992). La formación inicial del profesorado de educación secundaria: Papel de las didcticas especiales. En J. A. Meneses Villagra. & F. Lara Ortega (Comp.), Actas del I Congreso Nacional sobre formación del profesorado. "Modelos formativos del futuro profesor de enseñanza secundaria". E.U. del Profesorado de E.G.B. de Burgos, Universidad de Valladolid.

Garcia-Varcercel, A., & Fernandez Benitez, V. (1992). De la teoría a la párctica en la formación inicial del profesorado de Secundaria. En Meneses, ibidem.

Gene, A., & Gil, D. (1987). Tres principios básicos en el diseño de la formación del profesorado. Andecha pedagógica, 18, págs. 28–30.

Imbernon, F. (1989). La formación permanente del profesorado. Laia, Barcelona.

Krüger, M., & Veugelers, W. (1988). Recent developments in teacher activities and teacher education in the Netherlands. Reflexiones sobre la formación del profesorado. I.C.E. Universidad Autónoma de Madrid.

Mc Dermott, L. (1990). A perspective on teachers preparation in physics and other sciencies: The need for special courses for teachers. American Journal of Physics, 58, (8), págs. 734–742.

Perez Gomez, A. (1987). El pensamiento del profesor, vínculo entre la teoria y la práctica. Revista de Educación, 284, págs. 199–221.

Stenhouse, L. (1984). Investigación y desarrollo del curriculum. Morata, Madrid.

Villar Angulo, L. M. (1980). Enfoque modular de la enseganza, II. Aprender a enseñar, Cincel–Kapelusz, Madrid.

BEST COPY AVAILABLE

COUNSELING IN SECONDARY EDUCATION AFTER THE IMPLEMENTATION OF THE COMMON CURRICULUM

Cees Vermeulen, Klaas Bos, & Lidwien Cremers, Centre for Applied Research in Education (OCTO), University of Twente, Enschede, The Netherlands

Abstract

The impending introduction of a common curriculum in the first years of secondary education will make other demands on the school's system of counselling, especially on counselling with regard to choice of academic level and subjects, and will therefore compel schools to make changes to their system of counselling. The most appropriate system, its content and organization, was deduced in a study of eight schools with a highly developed system of counselling. By comparing these schools with matching schools with less developed systems found in a large scale survey of secondary schools, routes of improvement could be formulated by which schools could adapt their present system of counselling to the introduction of the common curriculum.

Introduction

In august 1993 a common curriculum will be introduced in the first three years of secondary education in the Netherlands (students 12-15 years old). While retaining the existing educational system of hierarchical ordered types of schools, the common curriculum is aimed at increasing the general level of knowledge and ability, as well at delaying the moment at which the choice for the future study or profession has to be made. This government initiated innovation represents a major upheaval of the current practice in secondary education. Amongst the elements that will have to change is the current system of counselling. In this study the necessary changes to this system were studied, particularly with regard to counselling children in the choice of school type and subject specialization. The aim of this study is twofold. First we will establish what criteria a counselling system has to meet in order to answer to the demands presented by the introduction of the common curriculum. In other words, the characteristics of a counselling system adapted to the new situation will be formulated. Next we will determine what changes will have to be made in the present situation in order to arrive at this system.

892

Method

In first instance a study was made of six schools with an exemplary counselling system. These schools were identified by the educational advisory centre commissioning this research. These schools were exemplary in the sense that these schools have a long tradition of giving a high priority to and putting a lot of effort in their system of counselling and have therefore acquired a commanding expertise and knowledge in this field. In these schools interviews were held with the school principal and with the counsellors, in order to describe the way the counselling was conducted in these schools. More specifically the aim was to determine the relationship between a number of school characteristics, like overall organization of the school, school type composition and size, and the system of counselling these schools had implemented. On the basis of these case-studies and on the basis of the constraints embodied by the common curriculum, a concept was formulated of a counselling system which can be regarded as the system, secondary schools should aspire to after the introduction of the common curriculum.

Simultaneously, a national survey was conducted in order to establish the present content and organization of the counselling in the first years of secondary education. A sample of 382 schools of the population of the about 1800 secondary schools in the Netherlands was drawn. These schools received a questionnaire directed at their system of counselling in general and at the counselling with regard to the choice of study and profession in special. On the basis of the responses provided by the 120 schools that filled in the questionnaire, we were able to compose an general picture of the counselling systems presently utilised in secondary schools. It transpired that for most schools there is still a long way to go before they arrive at the above described 'ideal' system of counselling.

The information supplied by the responding schools enabled us furthermore to select six schools which matched the exemplary schools we studied. The schools were matched on the basis of aspects such as school type composition and school size. These six match schools were the subject of another set of case-studies. Since these schools were part of the national survey, we already knew a number of global data concerning the schools organization and their system of counselling. In the interviews we held at these schools we could therefore restrict ourselves to gathering information specifically aimed at enabling the comparison with the exemplary schools.

This comparison has led to the identification of a number of factors which, to a large extent, can account for the differences in counselling systems existing between exemplary and match schools. Knowledge of these factors allowed us to describe the most appropriate routes a school will have to follow in order to adapt their present system of counselling to the system delineated in the course of this study as being most suited for the common curriculum. These routes obviously are pertinent to the specific organizational structure of the school.

Results

The study of the exemplary schools has resulted in a generalized concept of a system of counselling best adapted to the requirements posed by the introduction of the common curriculum. The main characteristic of such a system is a division of labour between the co-ordination of the counselling and the counselling itself. The

content of the counselling with regard to school and subject choice is decided on at the middle management level of the school (by the careers counsellor), while the actual work with the pupils is done by the teacher-counsellors. Another characteristic of exemplary schools is that they perceive counselling as a constant process rather than as a product with which the pupils are confronted on a limited number of occasions. These schools also maintain a high level of expertise in counselling, mainly by a strongly marked dedication towards in-service training. In the match schools less prominence is given to the counselling of pupils. In most instances only one person is responsible for developing as well as executing the careers counselling.

Conclusion

This study has produced a set of criteria a counselling system has to meet in order to adapt to the changes in secondary education, brought about by the introduction of the common curriculum in the Netherlands. Furthermore the data gathered has allowed us to pin-point routes of improvement for schools, the counselling systems of which do not meet these criteria at the present moment.

894

THE EVOLUTION OF WORK SATISFACTION, WORK LOAD, BURNOUT AND FEELINGS OF A SEVERE JOB CRISIS AMONGST DUTCH TEACHERS

Mart-Jan de Jong, Faculty of Social Sciences/RISBO, Erasmus Universiteit Rotterdam, Rotterdam, The Netherlands

Summary

In 1989 we have studied a national sample of secondary school teachers in the Netherlands. From this sample of almost 800 teachers 474 did take part in the study by answering a questionnaire. The response rate is nearly 60 per cent. The teachers have answered numerous scale items and questions about work satisfaction, demotivation and burnout, subjective workload and perceptions of a job crisis.

Work satisfaction could be compared with the outcomes of earlier studies, because we did use the same scale items. This comparison shows that the teachers satisfaction with his job has declined in the first half of the eighties, and stabilized in the second half of the eighties at this rather low level.

A historical comparison of the scores on the perception of the workload is more difficult, since former studies showed huge differences in approach. Study designs were different, and so were the type of questions or the ways in which the outcomes were presented. From the scarce possibilities to make comparisons, strictly on the basis of a few more or less identical questions, it shows that the perception of the work load has not undergone a significant change during the eighties. During this whole decade most teachers have considered their profession as a heavy and demanding job.

Our study into the evolution of burnout and feelings of a job crisis had to be different from former Dutch studies, since these earlier studies were unreliable. Also the construct validity of the scales could be questioned. Therefore we constructed various new scale items. These new items lead to much more reliable scales, and, we think, also to a more valid measurement of these phenomena. All this has important consequences for the conclusions based on this type of research. The alarming conclusions that were reported to the media earlier on the basis of methodologically weak measurements have to be qualified. Our approach shows much less grounds for panic about burnout and feelings of job crisis. This is due to the fact that our questions have been formulated in less vague way and seem to present a much better operationalisation of the concepts under study

here. This can be sustained by greatly improved outcomes of the factoranalysis and reliability analysis.
It does not seem to be true that one out every three teachers experiences feelings of burnout, as has been reported by Van Ginkel, but one in every six or seven. Our study shows that about five per cent of the secondary school teachers could be considered to be a severe case of burn out. His motivation has reached an absolute bottom level. Of course, this outcome does not give us much comfort either.
The proportion of teachers that is burdened by a strong feeling of a job crisis is estimated to be one and a half per cent. This too is much lower than has been reported by Prick. This means that every school with 70 teachers or more has on average at least one teacher that should quit his job as soon as possible, because his perception of the situation seems to be unbearable.
Finally, we want to emphasize once more that the difference between our conclusions and that of former studies does not seem to be consequence of a change in the educational climate, but is due to a more reliable a more valid method of measurement.

896

A METHODOLOGY FOR ELICITING DIFFERENT TYPES OF CONCEPT MAPS

N. Siregar, Bandung Institute of Education and Teaching, IKIP, Bandung, West-Java, Indonesia

Theoretical Background

Studies in teaching have not gained much benefits from the works of concept-mapping research. Because of its traditional orientation to discipline knowledge, concept mapping research has eluded the essential insight of its *psycholinguistic* basis. Experiences show that loosing such an insight cripples the studies in teaching aiming at, for instance, explicating teaching structures. Most teachers after completing their concept-maps were not too sure whether the maps reflect the way they teach or, rather, the familiar discipline knowledge (Fawns, 1985).

Ron Hoz's recent work *Didactic Use of Concept Mapping* (in Mahler, Hoz, Fischil, Tovly, and Lernau, 1991) can be seen as one of the initial attempts to break the tradition. The work carefully placed equal emphasis on the proposition link by increasing the sizes and categories of such links. This development interests anyone who tries to make *intelligible* use of concept-maps in the studies of teaching structures. More importantly, it provides supports to the depictions of actions as the proposition link component. The term *intelligible*, however, is emphasized to suggest that despite this important step, Ron Hoz has not touched the real problem of finding the essential insight of the psycholinguistic theory to guide the effort.

If studies of concept-mapping are to make more meaningful pedagogical contributions, they should recognize the distinction made in psycholinguistics between knowledge structures and macro- structures. Whereas knowledge structures correspond closely to the structures of discipline knowledge, macrostructures correspond to the abstracted structures of performances. These two structures, however, are closely related in terms of their production phases. Macrostructure derivations are the results of high order procedure (e.g. pedagogical) actions on the knowledge structures. Although in practice these two structures usually but unnecessarily undergo fusion, their individual surface characteristics persist. It can be shown that the top structure (see van Dijk and Kintsch, 1984) is predominantly of knowledge structures.

Methodology

Discussion on the methodology for eliciting subject matter knowledge structures and lesson macrostructures was undertaken using a school chemistry topic, namely, *Oxidation Reduction Reactions*. The participants for the study were 3 experienced teachers and one expert (PhD) of chemistry. Because of space limitation, the discussion provided is only a sketch of the work actually undertaken. For a complete report see Siregar's (1990) study.

A multi-method procedure was used for comparing different concept-maps which comprised the sorting, grid evaluation, and concept-mapping tasks. The first two tasks were each used for delineating and segmenting the subject matter concept-maps produced in the third task. In the first task, the subjects were presented with 16 small cards -- each with one concept written on it. They were asked to sort those cards and assign a label to each resulted pile. The results were that each participant formed a set of three piles and accordingly assigned a set of 3 labels. In the second task, a grid was presented to each participant with rows comprised the 16 concepts (i.e. the elements) and the columns comprised the labels (i.e. the constructs). The task was to assess the interrelationships between the elements and the constructs using numbers varying from 1 (strongly related) to 5 (not related). Each participant grid was appended to the expert's grid -- working between subject comparison. This matrix became the proximity matrix which served as input data for a multidimensional scaling procedure. Upon execution, the participant's individual constructs, except the third construct, merged into a neighbourhood; each conformed a distinctive region: Region I *Redox Process* and Region II *Electrochemical cells*. The third constructs did not form a neighbourhood but emerged instead into different quadrants of the (MDS)concept-map. These two opposing plot characteristics constitute the basis for differentiating common constructs from distinctive constructs.

In the mapping tasks, each teacher was provided with large card board sheet for arranging the previous 16 small cards. The procedure imposed no restriction; the participants were free to arrange the concepts in a way they wish and to dismiss some of the concepts provided (or to add new ones). After satisfied with their arrangements, the cards were then glued accordingly to the card board. This procedure should provide enough opportunities for the participants to establish their individual *salient* concept-maps.

The same procedure was also applied to the elicitation of lesson concept-map but involved only one teacher -- within subject comparison. The concepts were chosen from his own classroom presentation dealing with one of his previous constructs. However, in performing the task, despite the effort made to maintain their similar nature with the previous 16 concepts, the teacher spontaneously transformed the concepts into propositions. The teacher, likewise, used 'extended' (because involved arguments) proposition links. Although these deflections were unanticipated, the map resulted surprisingly resembled closely a macrostructure.

Results

The claim that the nature of concept-maps corresponds to a knowledge structure appears to be meaningfully supported. Both the expert and the teachers show a meaningful correspondences in Region I and Region II of their individual common constructs. The emergence of the distinctive constructs, however, does not

necessarily overrule such claim, because each is specific to the individual participants' perspectives on the redox topic. Examining the labels given to the distinctive constructs and the proposition links involved indicates individual orientations. The expert's orientation *fundamental particles* corresponds to his academic discipline; likewise, the teachers' individual orientations *example of*, *application*, and *origin of redox* correspond to their respective pedagogical treatments.

Meaningful correspondences are also observed either between the (MDS)concept-maps and the concept-maps or between the (MDS)lesson concept-map and the lesson concept-map. Similar configurations of concepts around the individual constructs are found both in the (MDS)maps and the maps. Checking the configurations further with the evaluation grids, shows that elements having strong relations with a particular construct correspond to one particular region in the concept-map; and ties in relationships in the grid correspond to sharing of the elements by two (or more) constructs in the concept-maps.

Tempering the production conditions of the lesson concept-map appears to be an important methodical basis which leads to the finding of concept-map typologies. Orienting the lesson concept-map production to classroom performance is necessary to induce different production modes. The lesson concept-map resulted, therefore, becomes distinctive in terms of sizes and types of either the concept component or the link component. The concept component of a lesson concept-map are propositions, such as, *must have both oxidation and reduction*, *standard half-cell needed*, *actual physical set-up of a complete cell*, etc.. As a result of changes in the size and type the lesson conceptmap shows that the concept component is integrated into the link component. Consequently, there is no essential functional difference between the concept component and the link component.

Conclusions

Applying convergency criteria to the findings, which is possible by the use of multi-method, two basic structural characteristics are found to underlie the differences between a conceptmap and a lesson concept-map. The first deals with the extent of elaborations applied in the individual structures, and the second deals with the types of dimensions in the individual structures. Whereas, in the concept-map, elaborations apply only to the concept component, in the lesson concept-map, elaborations apply both to the concept component and the link component. The elaborations in the link component are specific to lesson concept-map and concerned with progressions to bring the knowledge structures to the goal of the lesson. The organization of the lesson concept-map therefore is defined by the elaboration dimension, which specifically accommodates the thematic elaborations, and the progression dimension, which controls the segmentations and the assignments of a new theme for each segment of the lesson.

References

Fawns, R. (1985). Background mapping of teachers' cognate perceptions of ASEPS units. Research in Science Education, 13, 203-212.

Mahler, S., Hoz, R., Fischl, D., Tovly, E., &, Lernau, O. (1991). Didactic use of concept mapping in higher education: Application in medical education. Instructional Science, 20, 25-47.

Siregar, N. (1990). The Structures of Teachers' Explanations: Theory Methodologies, and Analysis. Unpublished PhD Thesis. Queensland (Australia): Queensland University.

Dijk, T., van, & Kintsch, W. (1984). Strategies of Discourse comprehension. New York: Academic Press.

A MODEL OF HOW STUDENTS LEARN HOW TO TEACH AS THE MEANINGFUL ACQUISITION OF PROFESSIONAL ACTIVITY

Gareth Harvard, & Richard Dunne, School of Education, University of Exeter, United Kingdom

Our paper critically examines the nature, acquisition and development of competence in teaching and the notion of the reflective practioner (Schon, 1987). We argue that both these notions are inadequately defined and understood. Initial training institutions have, until recently, neglected the challenge to articulate and justify their own version of competence and to demonstrate how such a view can be practically implemented in teacher education programmes, especially its implications for new routes to qualified teacher status. Moreover, there has been a particular neglect in assuming that experience is sufficient to learn how to teach and that learning from experience is unproblematic. This has allowed, to go unchallenged, the assertion that since experience is a necessary ingredient in learning to teach, then more experience enables students to teach better.

We seriously question this assumption by advocating a view of competence that emphasises the role of self-evaluation in the acquisition of professional activity and which clarifies the complementary roles of several mentors involved in its deliberate development at different stages in the learning cycle. Our model explains students' professional learning by adopting Ryle's (1949) notion of 'intelligent practice'. Recent attempts to depict teaching as craft apprenticeship and behavioural competences fail to recognise Ryle's important distinction between 'habitual' and 'intelligent' practices. " It is the essence of habitual practices that one performance is a replica of its predecessors. It is the essence of intelligent practices that one performance is modified by its predecessors. The agent is still learning." (p.42). But what precisely is being modified and how do you know when something has been modified? In other words, what is the nature of the evidence that learning has occurred?

Our model is elaborated to include not only classroom performances but also student teachers' knowledge structures or 'schema', and intellectual processes. The model attempts to integrate these three major elements in students' professional learning in this way: it describes the nature of the actions and activity that typically characterise teaching and students' reactions to such experiences; and it refers also to the production of performance. In the model, intellectual processes, particularly perception and categorisation of classroom events, comparison and analysis of specific cases and deliberation and judgement in

901

evaluating performances, select from and organise experience into schemas which are the repository of experiences. Processes and schemas become powerful mediators of experience and determinants of action. Learning, in this model, consists of the progressive development and modification of schemas and how these are structured over time. Whatever their precise structure, the development of professionally relevant schemas represent a further dimension of the growth of teaching competence.

In this scheme students' progressive development over a range of competences in planning, enacting and evaluating teaching is represented as progress from simple levels of performance to more complex expressions and levels of competence. The data for determining progress is collected from two broad domains: the overt experiences of and performances in classrooms and the nature and quality of students' thinking about those experiences and performances. We take seriously the need to be explicit about what thinking and learning processes are involved in observing and interpreting teachers' practices to help us to explain how professionally relevant knowledge is acquired and organised. So, in collecting data, we focus upon how, to what degree and under what conditions these capacities can be developed in training.

In order to represent the complex activity of teaching we have identified nine teaching dimensions. For each dimension there is a set of explicit criteria showing what improvement looks like as students progress to the more complex levels of competence. The criteria are intended as a learning aid. The onus for using them is on the student. They are used to develop student teachers' capacities for self-evaluation and realistic goal-setting. But this is conducted with particular techniques and under certain conditions so that evidence of students' teaching is being continually monitored: it is evaluated and corroborated by different mentors at different stages in the learning cycle.

There are three sources of data from which the student teachers and mentors monitor and evaluate progress and three categories of mentors who corroborate with the student what has been achieved: first, there are deliberately designed teaching episodes planned and enacted with the class teacher; second, these events are documented as validated accounts which students bring to a supervisory conference with a teacher-tutor; third, a range of validated accounts and their accompanying 'zones of development', are discussed with a university tutor in a supervisory conference. The validated accounts are constructed from agendas and their annotations; these accounts are validated in the sense that they provide valid evidence as a basis for further learning. They are then critically examined in supervisory conferences where students offer robust descriptions of their actual teaching, refer to relevant teaching dimensions and assess what has been accomplished and what is now possible for future development, all of which is driven by constant reference and persistent attention to specific aspects of the teaching dimensions and their accompanying criteria.

In both supervisory conferences the student teacher and mentor agree on a form of words that set a realistic aspiration for continuing progress. This written contract - which we call the 'zone of development' - outlines how the student teachers' classroom practices should develop in specific dimensions. The conduct of the conference is designed to model the intellectual processes of deliberation and judgement as essential ingredients of self-evaluation.

Data has revealed evidence of students' professional learning in all three aspects of our model: classroom performances, intellectual processes and professionally relevant schema. An analysis of the transcripts of conferences provides evidence of how it is possible to elicit students' practical reasoning, in which they offer, not only more robust accounts of situations, but also reasons for particular actions and activity and an understanding of how these illuminate what they plan to do next. There is further evidence of students' professional learning in their written evaluations for which we provide specific criteria. Students' capacities and dispositions for this type of self- evaluation as a particular form of appraisal, have already been well-rehearsed in their practical teaching with class teachers and in a more analytical fashion in both supervisory conferences. The criteria for written evaluation stress the importance of the collection and interpretation of data and the development of constructive, practical consequences, from analysis. Students are able to monitor their progress from descriptive but robust accounts of situations, to some tentative, justifiable analyses including children's reactions, alternative analyses in terms of tasks and resources, to justifiable explanations of children's responses to work and how these can be used to explain in practicable ways how to plan for the next phase of work.

Our data suggest that our stage-by-stage model of students' professional learning begins with a recognition and practical understanding of the nature of teaching actions and activity through coaching, modelling and practice; it continues with persistent attention to a representation of that teaching in the form of annotated agendas, and the manner in which structured discourse can successfully elicit students' practical reasoning in conferences, so that students' capacities for self-evaluation are being developed. The use of the structured framework of teaching dimensions and criteria is essential: they provide a basis for two important thinking procedures - analysis and reconstruction. The process by which conversation with the class teacher and the more formal discourse of conferences promotes the growth of thought is called dialectic. Our evidence is beginning to show that there is an analogous process in writing but it is one that student teachers find more difficult to develop and sustain. Both are effective in enabling partners with differing perspectives to reach mutually agreeable positions and in the process advance beyond the level of understanding that either partner possessed at the beginning. The dialectic is powerfully promoted when students are expected to devise and then persistently attend to various representations of the actions and activity of teaching against criteria as well as to test the impact of the criteria on the various products of and situations in which their teaching occurs.

In our view, these processes, techniques and conditions constitute the essence of 'intelligent practice' and the manner of students' professional learning as the meaningful acquisition of the professional activity of teaching. In response to the concern about the quality of initial teacher training and in recognition of the lack of knowledge related to the conditions that promote and sustain it, we have adopted a research perspective that emphasises the importance of seeing classroom performance in the context of student teachers' mental lives, specifically, the development of mental action, and the need to develop their capacities to process and organise professionally relevant knowledge and to use this organisation for designing, enacting and evaluating their professional practices and monitoring their further professional development.

DEVELOPMENTS IN STRATEGIC READING INSTRUCTION: EFFECTS OF A TRAINING OF TEACHERS

P. Reitsma, & S.E. Sliepen, Paedologisch Instituut Amsterdam, Duivendrecht, The Netherlands

In many research studies it has been shown that there is a relatively large number of children with serious difficulties in attaining fluent reading skills. Also, it has been firmly established that the skill of decoding single words is of crucial importance in the development of reading. The ability to decontextualize language, and specifically the ability to see words as entities that have phoneme sequences, has a significant effect on learning the spelling-sound correspondences and learning to identify written words. Children with difficulties in learning to read often appear to be handicapped in word decoding skills and a lot of instructional and remediation research has been aimed to find successful ways of improving word identification skills.

Word identification is not all that is involved in reading, of course. Having decoded the words, the reader must then comprehend their meaning and integrate the meanings available in sentences and paragraphs into coherent mental structures. The ability to comprehend written text requires complex cognitive processing strategies. Reading comprehension is not a matter of merely extracting meaning from print, but an active process in which the reader constructs meaning through a multitude of interactions with a text (e.g. Sliepen, 1991).

Therefore, there can be little doubt that if teachers wish to improve their pupil's reading comprehension they should directly teach the flexible use of various strategies in the reading of text. Yet, paradoxically, recent observational research has shown that teachers have been concerned almost exclusively with the end product of reading a passage as determined by means of a series of questions after reading (e.g. Durkin, 1979; Aarnoutse, 1991). For example, Durkin (1979) found that less than 1 percent of instructional time was spent on direct comprehension instruction. For most teachers comprehension instruction involves giving lots of assignments for students to do, asking students questions about assignments and focusing on getting the right answer.

In the present study a first question is whether teachers in special education classes for learning disabled children demonstrate the same pattern of teaching practices in reading comprehension lessons as found in regular educational classrooms. One could argue that in these classes direct comprehension instruction occurs much more frequently in daily practice. Learning disabled children are in great need for explicit instruction in almost all subject areas of basic education. The teachers who are specially trained for teaching learning disabled

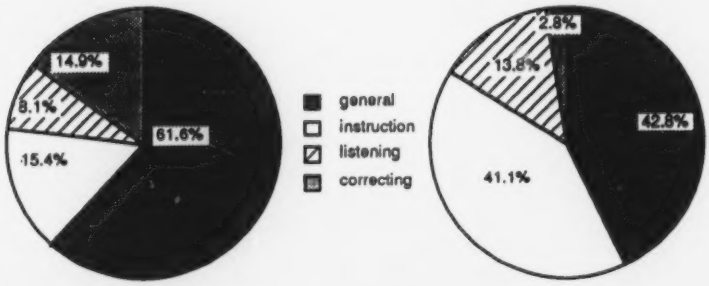
children are therefore used to deliberately and carefully prepare their pupils for all the steps necessary in schoolwork and to coach them very closely during the process of working through their schoolwork. Thus, relative to regular education teachers one could expect explicit instruction in reading comprehension much less rare in teachers of low ability readers.
Of course, a question remains how much time is actually spent on direct comprehension instruction within the constraints of a regular classroom situation. Given the earlier research (cf. Durkin, 1979) there is no reason to be overly optimistic about a high rate of direct instructional activities focused on reading comprehension strategies. The second question of this research therefore is whether it is possible to change the instructional behaviour of the teachers of low ability groups by training them to explicitly teach strategic reading skills. Earlier research (e.g., Duffy, e.a., 1986) suggested that it is well possible to train teachers to become more explicit in explaining how to use reading skills as strategies for comprehension. But this conclusion was based only on rather global rating scales of lesson transcripts and no direct observational data as for example obtained by Durkin (1979) were gathered to support this outcome.

Study 1

Method

In the first study we observed and video-taped each classroom teacher of a group of five during five reading lessons in order to obtain data with respect to the practice of teaching reading. Both an interval time-sampling observation method and a rating scale on quality of instruction was used. The observation system was adapted to the earlier work of Durkin (1979) and Aarnoutse (1991) and consisted 14 coding categories to describe the most important activities in the classrooms. The scheme included categories such as general management, attending to and/or correcting decoding, direct instruction in generating predictions on basis of the text, focusing on getting the right answer, etc.. Interobserver reliability appeared to be quite acceptable (.78) and intra-observer consistency was about .82. The rating scale was adapted from the work of Duffy e.a. (1986) and consisted of 11 categories concerning the information the teachers presented about reading comprehension strategies, the means by which this information was conveyed (e.g. use of modeling, review, or feedback), and the integration within the lesson of with other lessons. Interobserver reliability appeared to be .93 and intra-observer reliability was about .95.
Then a series of five intensive workshop sessions was provided to the teachers during which relevant instructions for reading comprehension were explained, discussed, modelled and practiced. Use was made of video-tapes of their own lessons. The teachers were taught how to integrate explicit strategy instruction into their on-going reading lessons and were guided to develop their own instructional plans. The training sessions lasted about 2,5 hour each and were scheduled once a week.
Three months after the workshops were finished another series of observation and rating of reading lessons was conducted. The procedure and scales were similar to the ones used in the pre-training observations.

Results
The results of the time-sampled observations are summarized in the piecharts below.



The results show a substantial decrease in the time devoted to general management procedures and product-oriented guiding behaviour (62% to 43%). At the same time a significant increase was observed in the time spent on direct instruction of reading strategies. In fact almost twice as much time was devoted to explicit instruction and guidance in the use of strategies (15% to 41%).
The results for the rating scale for teachers explicitness showed a similar pattern. There was a statistical significant increase of 7.0 to 23.8 in rating data indicating that an impressive improvement in quality of instruction was obtained.
Thus the data obtained from this first study seem to indicate that it is well possible to change actual teaching behaviour in the classrooms by means of a series of workshop sessions.

Study 2
Although the results of the first study were quite positive, a few reservations should be made. First the number of teachers cooperating in this study was rather limited and prevents firm statements with regard to generalization. Second, because of the frequent visits to the classrooms in order to observe the daily reading lessons, it might be that teachers became aware of the focus of the observations (they were not told about the objectives though) and changed and improved their teaching practices accordingly.

Method
A second study was therefore designed in order to extend and replicate the findings of the first study. Two improvements in the design were established. First, the number of teachers cooperating in the research was increased. Second, a control group of teachers not receiving a training was added. The study involved 18 teachers, 10 receiving a series of training workshops and 8 serving as a control group. Because the observation results in study 1 corresponded very closely to the rating data and because the procedure of time-sampled observation was very

time-consuming with respect to both the collection and the processing of data, it was decided not to use time-sampled observation data in this study.

Results

Preliminary analyses of the data corroborated the findings of the first study. The averaged pretraining rating score was 9.1 and no difference was found at that time between the group of teachers involved in the later training sessions and the control group of teachers. Preliminary analyses of the posttraining observations and ratings do suggest an important improvement in rating scores after the training sessions in the experimental group. While the control group also seems to improve their instructional behaviour, the group of teachers that received the experimental training were significantly scored as demonstrating qualitative better instruction in reading comprehension strategies.

Conclusion

Observations of actual teaching behaviour indicates that teachers do not spend much time on direct instruction of reading comprehension strategies. There is no reason to assume that teachers in schools for special education do differ in this respect to their colleagues from regular elementary education. Also, although not much time is devoted to actual comprehension instruction, rating scores also show that the instructional activities are to be considered relatively poor according to standards derived from recent accumulated knowledge about effective teaching of strategies.

However, it seems well possible to train teachers to actually incorporate new approaches toward reading comprehension strategies into their on-going instructional behaviour in the classrooms. Both the observational time-sampled data and the observational rating data strongly suggest that teachers can improve the teaching of reading comprehension strategies. A series of several intensive training workshops for the teachers appeared to significantly influence their instructional activities. An important question for future research is of course whether these changes have any impact on the comprehension performance of their pupils.

References

Aarnoutse, C. A. J. (1991). Begrijpend lezen in het basisonderwijs. In P. Reitsma & M. Walraven (Eds.), Instructie in begrijpend lezen. Delft: Eburon.

Duffy, G. G., Roehler, L. R., Meloth, M. S., Vavrus, L. G., Book, C., Putnam, J., & Wesselman, R. (1986). The relationship between explicit verbal explanations during reading skill instruction and student awareness and achievement: A study of reading teacher effects. Reading Research Quarterly, 21, 237-252.

Durkin, D. (1979). What classroom observations reveal about reading comprehension instruction. Reading Research Quarterly, 14, 481-533.

Sliepen, S. E. (1991). Tekstbegrip bij zwakke lezers - I / II, JSW, 76, 2, 24-27 & 3, 18-20.

THE CONTRIBUTION OF JOURNAL WRITING TO STUDENT TEACHERS' PROFESSIONAL LEARNING

Chris James, & James Calderhead, School of Education, University of Bath, United Kingdom

Journals, diaries and records of experience have become used increasingly in initial teacher education courses as a means of promoting awareness, reflection and professional growth, and profiling the development of teachers' competence. Whilst there is considerable enthusiasm for their use (e.g. Holly & McLoughlin, 1989; Handal & Lauvas, 1987) little is known about how they aid student teachers' professional learning. This paper reports an enquiry into the contribution of one such device to the professional development of a group of postgraduate student secondary teachers. The enquiry aims to explore what journal writing entails for student teachers and to describe the processes by which journal writing might enhance their learning, setting this within a wider context of teachers' professional learning (e.g. Calderhead, 1988) and discussions of reflective practice (Clift, Houston & Pugach, 1990).

The Record of Student Experience (ROSE) is a structured journal which the students in one British university are asked to keep. It requires them to note learning experiences that they encounter both in the university and school components of their course, to state what it is that is valuable about these experiences, to use them to formulate an agenda for their further professional development, and to discuss their experiences and their evaluations with supervising teachers and university tutors.

A sample of 16 students were followed through their one year course and were interviewed on four separate occasions throughout the year. The interviews focused on students' general reactions to the ROSE, on the nature of the experiences logged in their ROSE journals, and to the learning processes involved in recording, reflecting on and discussing particular teaching experiences.

The interview transcripts have been analysed to address the following questions:

- What are the students' reactions to the task of recording and analysing their learning experiences?
- How do students learn from this task?
- What do they learn from the task?
- Do all students benefit from the task? If not, what characterises the students who learn successfully?
- What is the relative contribution of the recording, analysing and discussion components of the task?

Our initial analysis of the data suggests that students vary considerably in their response to the journal writing task. Some value it highly, others see no value in it at all for their professional development. The former appear already to adopt a reflective analytical approach to their everyday life experience; the latter appear oriented to learn in other ways (e.g. modelling themselves on a 'good' teacher). Over the course of the year, some students make a transition from the latter to the former, changing their ideas and attitudes towards the process of learning to teach. The analysis of interview transcripts in relation to particular learning experiences suggests that the ROSE may serve several functions in professional development.

- Writing itself appears to make students more aware of the complexities of teaching and of the processes of learning to teach; the process of writing about their experience heightens students' awareness;
- A written record of their experience appears to enable students to take an objective stance on their experience, looking at it more critically and adopting a problem-solving approach;
- Writing about their experience appears to lead student teachers to value their experience, to develop a greater sense of ownership of their own professional development and to develop a sense of responsibility for their own professional learning;
- Writing about experience seems occasionally to have a cathartic effect, enabling student teachers to confront and analyse anxiety-provoking incidents;
- A written record of experience enables an agenda to be easily formed for discussion with others about practice;
- A record of their experience over time appears to give student teachers an appreciation of their own professional development, and a level of confidence in their own developing abilities.

Examples of each of these functions will be given. Issues will be raised about our understanding of the processes of student teachers' professional learning, and of the possible contribution of journal writing activities in initial teacher education.

References.

Calderhead, J. (Ed.) (1988). Teachers' Professional Learning. Lewes, Sussex: Falmer Press.

Clift, R. T., Houston, W. R., & Pugach, M. C. (1990). Encouraging Reflective Practice in Education. New York: Teachers' College Press.

Handal, G., & Lauvas, P. (1987). Promoting Reflective Teaching. London: Open University Press.

Holly, M. L., & McLoughlin, C. (1989). Professional Development and Journal Writing. In M. L. Holly & C. McLoughlin (Eds.), Perspectives on Teacher Professional Development. Lewes, Sussex: Falmer Press.

THE TEACHING PROFESSION: TEACHERS' REPRESENTATIONS OF THEIR POSITION IN SOCIAL SPACE

A. de Lima, University of the Azores, Portugal

The notion of social representations (Moscovici, 1961) helps us to understand the way actors construct the social reality in which they move (Gilly, 1989).
Boudieu (1979) has stressed that the social agents produce, not only classifiable acts, but also classification acts. Berger an Luckman (1985) have demonstrated that society is a twofold reality, objective and subjective.
Therefore, we find it important to the systems of symbolic constructions that individuals build, in order to decodify the complexity of the world that surrounds them.
Our study focus on teachers' representations, especially on those which are an expression of their particular ways of percepting their position in the social space.
In Portugal, the only data to once knowledge that can be included in this discussion are the proposed by the Comissao 14/Me/88 (1988: 1205) about teachers' self-appreciation of their social prestige. These data suggest that the "positive" self-appreciation is mentioned by an important proportion of teachers, but that those who see themselves in an "intermediate position", in terms of prestige are numerically more important.
Our data do not refer to an appreciation of the social prestige that teachers think they enjoy in the social system, but rather to their representations of the .relative position occupied by the main categories of teachers' primary, secondary, and universitary) in the social space. they do not refer to the way they think other people see them, but rather to the way they see themselves, and their colleagues, in terms of social standing.
We studied the variations of the representation (measured by a 9-point likert-type scale) in association with a wide set of factors, grouped as follows: personal characterization variables (sex), professional characterization variables (occupation prior to teaching, academic subject, level of prestige of the subject matters that are taught, and job situation), social class origins, school capital (academic habilitations, and average school grade), relation to the school institution (reason for choice of the highest studies accomplished) and relation to teaching (prior commitment to a teaching job, reasons for deciding to teach, job satisfaction, and propension for quitting teaching).

For the purpose, we distributed written questionnaire to all teachers working in the secondary schools of the Autonomous Region of the Azores, Portugal. A response rate of 40.p % was obtained. The data were statistically treated in terms of significance tests, analysis of variance and correlation analysis.

The results showed three major conclusions:

- secondary teachers' representations of the relative positions occupied by the main categories of teachers in the social space are very homogeneous where refering to the social position of university teachers, regardless of the differentiation factor under consideration;
- representation concerning their own social position show a slight variation in some aspects, but these are very small, and never reach statistical significance at he 0,05 level;
- representations refering to primary teachers show important variations, practically in association with sex, job situation and propension for quitting teaching.

Our study suggests that this might be an important feature of secondary teachers' professional culture.

THE STUDY OF TEACHERS' PROFESSIONAL BIOGRAPHIES: METHODOLOGICAL EVALUATION

Geert Kelchtermans, & Roland Vandenberghe, Centre for Educational Policy and Innovation, Catholic University of Leuven, Leuven, Belgium

Abstract

The paper presents a methodological reflection on the research on teachers from the "biographical perspective". First the research procedure, as developed in a research project on teachers professional biographies, is described and argumented in detail. Further a series of methodological questions, inherent to this kind of research, are discussed: e.g. the consequences of the interactive nature of the research situation and the consequences for the nature of the research data; the functioning of autobiographical memory; the constructive character of the autobiographical selfthematisation by the respondent; the narrativity in the research data; reliable and valid interpretation; generalisation; researchers' subjectivity etc..

The final conclusion of the methodological evaluation is that the biographical perspective forms an important approach for the construction of a grounded and encompassing theory of teachers' professional behaviour and its development.

Stating the problem

Since the end of the seventies educational researchers are interested in studying the development of teachers' professional behaviour from the so called "biographical perspective". This interest can be considered as the educational branch at the old trunck of the biographical method in sociology (Chicago School), that also has branches in psychology and in historiography (e.g. the oral history–movement) (Kelchtermans & Vandenberghe, 1990).

In most publications however, the specific procedures for data collection and –analysis remain implicit. In some cases one can raise serious questions about the methodological quality of these (qualitative) studies. In the paper we first present a grounded description of a research procedure, as developed for a study from the biographical perspective. On the basis of this description a number of methodological issues, concerning the biographical approach in educational research, are discussed. Specific solutions to avoid biases are described (Kelchtermans et al., 1991).

A research procedure for the study of teachers' professional biographies
The researchproject aimed to reconstruct the career experiences of primary school teachers. By means of 'stimulated autobiographical selfthematisation' we reconstructed the career, focussing on the subjective meaning. Through this reconstruction we gained insight in the professional development, defined in terms of professional self and subjective educational theory.
The final research procedure, which we developed through several pilot studies, consisted of a cycle of biographical interviews, completed with questionnaire data, observations in school and classrooms, interviews with other informants (e.g. the school leader). Data–collection and analysis were interwoven throughout the procedure in a cyclical way. The complete procedure was described and argumented in an extended "research scenario". In the paper the content and function of this research scenario are extensively discussed (Kelchtermans, in press).

Methodological evaluation
Related to the research procedure, a number of methodological issues and problems, inherent at the biographical method are amply discussed.
First we discuss the nature of the research data: we explore systematically the interactive nature of the research situation and the consequences for the research data. The respondents were stimulated by the researcher to retrospectivally reconstruct their career, from their personal interpretation of the experiences. Retrospection leads to the problem of (autobiographical) memory.
Further the reconstructive nature of the data implies that the respondent himself actively construes a coherent career story and so affects the story. Because of the social, interactive research situation it is also important to pay attention to the interaction between researcher and respondent. The way the respondent perceives the researcher (and vice versa) influences the nature and quality of the produced data. Also the researcher's attitude and behaviour during the interview have an impact.
Another important issue is the narrative quality of the research data. In the reconstruction we gave maximal space to the spontaneous language of the teachers ("the teachers' voice"). The linguistic nature of the research data therefore is narrative (anecdotes, stories, metaphors), which has implications for the analysis.
Apart from the data collection also the analysis and interpretation of these data are dealt with in the paper. We discuss the quality norms (reliability, validity), the possibilities for generalization and also pay special attention to the question of controling the researchers' subjectivity in this kind of research.

Conclusion
In a final paragraph we summarize the possibilities and difficulties of the biographical perspective in educational research are discussed. Our conclusion is that the biographical perspective (as a variant of other qualitative methods) forms an important approach to develop grounded and encompassing theories of teachers' professional behaviour and the way it develops. Some of our conclusies and concrete procedural suggestions in the paper however even go beyond the

biographical perspective and are important for qualitative educational research in general.

References

Denzin, N. (1970). The Research Act. A Theoretical Introduction to Sociological Methods. Chicago: Aldine.

Denzin, N. (1989). Interpretive Biography. Newbury Park: Sage.

Kelchtermans, G. (in press). Teachers and Their Career Story. A Biographical Perspective on Professional Development.. In C. Day, J. Calderhead, & P. Denicolo. Research on Teacher Thinking: Towards Understanding Professional Development. London: Falmer Press.

Kelchtermans, G., & Vandenberghe, R. (1990). Het biografisch perspectief in onderzoek van het professioneel handelen van leerkrachten. Pedagogisch Tijdschrift, 15(3), 126-139.

Kelchtermans, G., Vandenberghe, R., Jaeger, A., de, & Indenkleef, M. (1991). De professionele ontwikkeling van leerkrachten vanuit het biografisch perspectief. Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs, 7(5), 284-302.

Plummer, K. (1983). Documents of Life.An Introduction to the Problems and Literature of an Humanistic Method. London: Allen Unwin.

Rubin, D. (Ed.). (1986). Autobiographical Memory. Cambridge: University Press.

THE PROFESSIONAL DEVELOPMENT OF PRIMARY SCHOOL TEACHERS

Mieke Clement, & Roland Vandenberghe, Centre for Educational Policy and Innovation, Catholic University of Leuven, Leuven, Belgium

Abstract

The professional development of teachers is often considered as a central factor in establishing a local school policy in general and an improvement program in particular. Nevertheless, in most studies the definition of professional development is vague. In our study we conceptualize professional development as a sequence of learning experiences; each experience exists of an opportunity to learn and the individual action of the teacher towards this opportunity. Relying on the results of 45 in depth interviews and case–studies in 5 primary schools, we will elaborate on the individual and organizational components of the professional development. This study is part of a larger study on the relationship between individual characteristics, working conditions in a school and professional development.

Introduction to the problem

The professional development of teachers is often considered as a central factor in establishing a local school policy in general and an improvement program in particular. If one wants to improve the school, one should create optimal working conditions and opportunities for teachers to develop professionally. Indeed, individual professional development as well as favourable working conditions are believed to influence student achievement. A continuous professional development is a necessary condition for the implementation of effective instruction.

Up to now, however, 'professional development of teachers' is a concept which is not used uniformly in the literature. Some researchers distinguish several stages which characterize the career of teachers and which are often connected with more general developmental phases and tasks (Erickson, 1968; Levinson, 1978, Gould, 1978; Kohlberg, 1984; Loevinger, 1976; Kegan, 1982). Others use a biographical approach in order to outline the professional development (Kelchtermans & Vandenberghe, 1990; Kelchtermans et al., 1991). Still others work in a very pragmatic way. Rowe and Sykes (1989) e.g. determine the professional development of teachers by counting the number of in–service programs they attended.

Since professional development seems to be so important in view of the improvement of education and since it is not yet completely defined, nor analysed, we consider this concept as the central theme of our study.

Design and Method

We combine qualitative and quantitative research methods. First, we want to describe and conceptualize the way teachers themselves assess their individual professional development and the factors they point out as crucial determinants of this development. An inductive, interpretive methodology is most appropriate. So we started with 43 in depth interviews of 18 novice (less than 5 years of teaching experience) and 25 experienced elementary teachers (15 / 20 years of teaching experience). We expect that both categories will differ in their opinions about professional development. These interviews will enable us to analyse individual accounts, descriptions and interpretations of professional development. Since the construction of meaning is essentially an interactive, social process, 5 primary schools and the activities of the team will be studied for a longer period of time (5 case–studies beginning in the spring of 1992). This will allow us to judge more accurately the significance of the social (school)context for the professional development of teachers.

Relying on the results emerging from the interviews and the case–studies, it will be possible to generate an instrument to measure quantitatively the professional development of teachers in a larger, representative sample of primary schools. The results of this measurement will be linked to some characteristics of the school, such as the professional culture, collective sense of efficacy, innovative structures, to be measured with other already existing instruments.

Results

Based on a literature study, the notion of professional development was elaborated during the first stage of the research. We hypothesized that professional development can be characterized as a sequence of learning experiences, starting at the moment one enters training until one leaves the profession, voluntary or not. Each experience exists of an opportunity to learn offered by the school and the individual action of the teacher towards this opportunity.

We checked this preliminary conceptualization through the 43 in depth interviews with primary school teachers. Some interesting ideas and patterns emerge. The responses to the open–ended questions clearly indicate that teachers consider their own professional development as a continuous process. Professional development is for the larger part definitely linked with classroom activities. Some teachers specifically refer to instructional improvements. Others give an account of development concerning their relationship with their pupils. These results are clearly in line with the traditional image of teachers getting their professional satisfaction essentially from their involvement and interaction with pupils (Lortie, 1975). Most of the results also confirm Hoyle's (1975) description of the restricted professional, the teacher who is focusing primarily on teaching and the daily professional practice in the classroom. Often such restricted professionals act quite isolated and are only to a minimum degree interested in activities outside their classrooms. The interests of extended professionals on the other hand go far

beyond their own classroom. They focus more on being a member of the school team. They relate classroom activities to the policy and goals of the school. They stress the importance of professional collaboration and show also a more intensive involvement towards non–teaching activities.
In both categories teachers explain their opinion about professional development by referring to both personal and organizational factors. In a further analysis of the data professional development is related to individual characteristics, such as the subjective educational theories, the sense of efficacy and the career perceptions. As for the organizational features, we expect – inspired by Mc Laughlin and Yee (1988) – that certain qualities of the school context influence the occurrence of learning opportunities. We suppose that organizational characteristics such as the adequacy of resources, common goals, collegiality, a problem–solving orientation, an investment–centred reward structure are important organizational characteristics.

Conclusion

The results lead to a valid definition of professional development, and to a clear image how this development is related to some individual and organizational characteristics.

References

Erickson, E. H. (1968). Identity, youth and crisis. New York: W.W. Norton.

Gould, R. L. (1978). Transformation: Growth and change in adult life. New York: Simon & Schuster.

Hoyle, E. (1975). Leadership and decision–making in education. In M.G. Hughes et al. (Eds.), Managing education: The system and the institution. London: Holt, Rinehart & Winston.

Kegan, R. (1982). The evolving self: problems and process in human development. Cambridge, MA: Havard University Press.

Kelchtermans, G., & Vandenberghe, R. (1990). Het biografisch perspectief in onderzoek van het professioneel handelen van leerkrachten. Pedagogisch Tijdschrift, 15(3), 126–139.

Kelchtermans, G. et al. (1991). De professionele ontwikkeling van leerkrachten vanuit biografisch perspectief. Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs, 7(5), 284–302.

Kohlberg, L. (1984). The psychology of moral development. San Francisco: Harper & Row.

Levinson, D. et al. (1978). The seasons of a man's life. New York: Knopf.

Loevinger, J. (1976). Ego development: conceptions and theories. San Francisco: Jossey–Bass.

Lortie, D. C. (1975). Schoolteacher: A sociological study. Chicago: University of Chicago Press.

McLaughlin, M. W. & Yee, S. M. (1988). School as a place to have a career. In A. Lieberman (Ed.), Building a professional culture in schools. New York: Teachers College Press.

Rowe, K. J. & Sykes, J. (1989). The impact of professional development on teachers' self perceptions. Teaching and Teacher Education, 5(2), 129–141.

917

BEST COPY AVAILABLE

CAREER STORY, SELF AND SUBJECTIVE THEORY: A BIOGRAPHICAL APPROACH TO PROFESSIONAL DEVELOPMENT

Geert Kelchtermans, & Roland Vandenberghe, Centre for Educational Policy and Evaluation, Catholic University of Leuven, Leuven, Belgium

Abstract

The paper reports a study of teachers' professional development from the biographical perspective. The research question is: can we gain insight in the professional self and the subjective educational theory (as results of the professional development) and their evolution throughout the career, by reconstructing the professional biography of the teachers?

Ten experienced primary school teachers were studied by a research procedure – inspired by the "grounded theory approach"– consisting of a cycle of biographical interviews, completed with other qualitative methods, to reconstruct the teachers' career stories. Throughout the case studies of three teachers the developing theoretical framework is explicitated and illustrated.

In the conclusions the meaning of the research findings for teacher pre– and in–service training is discussed.

Conceptual framework and research questions

The notion "professional development" refers to the evolution of teachers' professional behaviour during their career (Holly & Mc Loughlin, 1989). The professional development is studied from a "biographical perspective" (see e.g. Hirsch et al., 1990; Huberman et al., 1989; Sikes, Measor & Woods, 1985). As as starting point we assume that professional development results from a continue interaction between the teacher and his professional environment. The development leads to a "professional self" and a "subjective educational theory". The conceptual framework was strongly influenced by the theory of symbolic interactionism.

Our specific research question is then: can we gain insight in the professional self and the subjective educational theory as they develop during the career, by reconstructing the professional biography of the teachers? More specifically we are also interested in the determinants of the development a conceptual framework (Kelchtermans, in press; Kelchtermans et al., 1991).

Research procedure

In the paper one study, as part of a larger project, is presented. Two pilot studies preceded this study. In these pilot studies we developed a conceptual framework as well as a concrete research procedure. Methodologically we were strongly inspired by the "grounded theory"–approach.

In the study 10 primary school teachers in their midcareer (between 15 and 25 years of classroom practice) were involved. The research procedure consists of a cycle of three biographical interviews, observations in the school and the classroom, completed with data from questionnaires (reconstruction of the formal career) and interviews with other informants (e.g. school leader). In the procedure datacollection and analysis are interwoven. The results of an interview or observation always builded the basis for the next interview or observation session. Through the interviews teachers were stimulated to "autobiographical selfthematisation": this is a retrospective reconstruction of their career as a teacher from their personal view. We emphasized thus the subjective meaning and personal interpretation of the experiences by the teacher. With "professional biography" we therefore mean the "career story" and not the formal career.

Finally the procedure of datacollection and –analysis resulted in a "synthesis text". This text (with a fixed structure for all respondents) contained a reconstruction of the career story and the interpretations on professional self, subjective educational theory en their determinants. This synthesistext was presented to the respondent for control.

Results

In the paper we discuss the most important results of the study. The notions professional self and subjective educational theory, developed in the pilot studies, are now more concretely illustrated (content) and their interrelations were stronger empirically grounded. Our conceptual framework thus developed from a cluster of sensitizing concepts to a grounded theoretical framework.

In the first place the narrative approach of teachers' careers resulted in "thick descriptions" of their development throughout their professional life. In these career stories certain persons, incidents and phases appeared to be 'turning points' or 'key experiences'. We refer to them as "critical persons, incidents and phases". They distinguish themselves from the innumerous other memories by the pregnance of their meaning. They had an impact on the professional self and the subjective educational theory (and thus indirectly also on the professional behaviour). In the paper we illustrate these notions by means of a case study of three respondents (colleagues from the same school). Throughout these case studies also the coherence between the different notions in our theoretical framework will become clear.

A second series of results are related to the content of the notions "professional self" and "subjective educational theory". Out of the analysis we construed several components, that make it possible to differentiate the context of the notion. E.g. in the data concerning "professional self" we distinguished six components: self–image, self–esteem, jobsatisfaction, motivation, future perspective and task perception. In the paper we give an overview of these components and illustrate their content with data from our research.

Discussion

In the final discussion we reflect upon the relevance of this research approach. Although the number of respondents in this intensive time– and workconsuming study was rather small, their stories help us to understand the reality of what it means "being a teacher". We look at the school reality through the eyes of the teachers. Their way of seeing things is of central importance to our work. This way we get new and unusual information about the teachers job, e.g. their job motivation (choice, teacher education, induction period, eventual demotivation and burnout). For teacher pre– and inservice education important lessons can be learned from this study. We therefore formulate several concrete suggestions on these topics. We finish our paper by sketching the contours of a new study in this project, which on the one hand focuses on teachers who leave education voluntarily. On the other hand we describe the first steps in developing a Career Questionnaire. This questionnaire, as a complementary research tool to the interviewprocedure, will make it possible to study the professional biographies of a larger group of teachers.

References

Hirch, G., Ganguillet, G., Trier, U. P., Egli, H.& Elmer, H.-R. (1990). Wege und Erfahrungen im Lehrerberuf: Eine lebensgeschichtliche Untersuchung Ober Einstellungen, Engagement und Belastung bei ZGrcher Oberstufelehrern. Bern: Haupt.

Holly, M. L., & MC Loughlin, C. (1989). Perspectives on Teacher Professional Development. London-New York-Philadelphia: Falmer.

Huberman, M. Grounauer, M., & Marti, J. (1989). La vie des enseignants. Evolution et bilan d'une profession. Neuchtel-Paris: Delachaux et Niestl.

Kelchtermans, G. (in press). Teachers and Their Career Story. A Biographical Perspective on Professional Development.. In C. Day, J. Calderhead, & P. Denicolo. Research on Teacher Thinking: Towards Understanding Professional Development. London: Falmer Press.

Kelchtermans, G., & Vandenberghe, R. (1990). Het biografisch perspectief in onderzoek van het professioneel handelen van leerkrachten. Pedagogisch Tijdschrift, 15(3), 126-139.

Kelchtermans, G., Vandenberghe, R., Jaeger, A., de, & Indenkleef, M. (1991). De professionele ontwikkeling van leerkrachten vanuit het biografisch perspectief. Nederlands Tijdschrift voor Opvoeding, Vorming en Onderwijs, 7(5), 284-302.

Nias, J. (1989b). Primary Teachers Talking. A Study of Teaching as Work. London-New York: Routledge.

Sikes P., Measor, L., & Woods, P. (1985). Teacher Careers. Crises and Continuities. (Issues in Education & Training Series, 5). London-Philadelphia: Falmer.

PRIMARY SCHOOL TEACHERS BETWEEN AUTONOMY AND COLLEGIALITY

Katrine Staessens, & Roland Vandenberghe, Centre for Educational Policy and Innovation, Catholic University of Leuven, Leuven, Belgium

Abstract

Many studies about professional relationships among teachers suggest converging results. Collegiality and collaboration are considered as necessary conditions for curriculum and school development. Opposed to the image of 'the collegial school', the autonomy of teachers is equated with professional isolation. Autonomy and collegiality are conceived of as two polar concepts. In the paper we will surpass this polarity in favour of a more circular thinking about the tension between autonomy and collegiality, as it is experienced by teachers. In the study 43 teachers in 13 primary schools were involved. Evidence was found for the existence of various forms of collegiality and autonomy. Each of them can have a positive or a negative meaning depending on the working conditions in the school.

Introduction of the problem

During the last decennia, issues related to professional relationships and collaboration among teachers have received a great deal of attention by practitioners as well as by researchers. Many studies suggest converging results, namely that there exists a relationship between the quality of the professional collaboration among teachers and for instance the effectiveness of the school and the implementation of innovations. Professional relationships are considered as an important instrument for curriculum and school development (see Berman & McLaughlin, 1978; Little, 1981; 1982; Wilson & Corbett, 1983). Within this perspective, the image of 'the collegial school' (Nias, Southworth & Yeomans, 1989; Rosenholtz, 1989) has become a normative concept. On the opposite, the autonomy of the individual teacher is mostly defined in terms of professional isolation (Hargreaves, 1990).

However, when looking more carefully at the results of various studies, it seems that only a small percentage of the schools corresponds to the normative image of the collegial school. The professional autonomy of the individual teacher persists as a fundamental characteristic of 'being-a-teacher', in spite of the measures that have been taken to stimulate collegiality (Little, 1990; Zahorik, 1987; Zielinski &

Hoy, 1983). In this perspective, 'autonomy' and 'collegiality' appear as two polar concepts.
In the paper, we present the results of the qualitative part of a larger study in which a conceptual framework was constructed, based on empirical evidence: the relationship between autonomy and collegiality is re–evaluated. In other words, a framework is developed in which meaningful collegiality as well as accepted and competent autonomy are empirically described and assessed.
In the paper that will be presented, we will answer the following two questions:

- What types of autonomy are compatible with professional collegiality?
- What variants of collegiality leave room for professional autonomy?

By asking these questions, we try to avoid a polar thinking between autonomy and collegiality in favour of a more circular thinking.

Design and Method

In the study 43 teachers (18 beginning and 25 experienced teachers) in 13 Primary Schools were involved. The interviews took place in January 1992. In order to obtain comparable information across several teachers and schools, a semi–structured interview–guide was constructed. The conversations with the teachers took on the average 45 minutes and were registered on tape.
Different transformations were carried out on the protocols. The interviews were coded by an a posteriori constructed coding scheme. Meaningful code–combinations were retrieved by a computer program (FYI–3000–Plus).
The qualitative analysis consisted of two phases. In the vertical analysis, we tried to understand the problem of the meaning of autonomy and collegiality and of the tension between them for each teacher individually. In order to understand the teacher's reactions and ideas about collegiality and autonomy and to construct a valid interpretation, information about the school and about the working conditions were also collected. In the horizontal analysis we looked across all the teachers in order to find patterns that can explain the field of tension between autonomy and collegiality.

Results

Using this qualitative method, it seems possible to discover the meanings teachers give to autonomy and collegiality and to the tension.
The results of the study present a more differentiated picture than the polar image we mentioned already. Evidence was found for the existence of various forms of collegiality, in which not only the form but also the content is important (e.g. storytelling; aid and assistance; sharing of methods and materials; collegial work). These forms differ to the extent in which they induce mutual interdepence. We could also discern different forms of autonomy or individualism (e.g. constrained individualism; strategic individualism; elective individualism; assigned individualism).
Each variant of autonomy and collegiality can have a positive or a negative meaning, depending on the working conditions in the school and the degree of professional development of the teacher. Some variants of collegiality preserve the position of isolation of the individual teacher. On the other hand, some variants of autonomy are a good breeding ground for collegiality.

In the paper we will also elaborate some differences between the beginning teachers and the experienced teachers.

Conclusion and discussion

It seems possible to determine empirically different variants of autonomy and collegiality. Further, various configurations can be distinguished concerning the combinations of variants of autonomy and collegiality. Finally, evidence was found for the relation with the schoolcontext.

References

Berman, P. & McLaughlin, M. (1978). Federal Programs Supporting Educational Change, Vol. VII. Factors Affecting Implementation and Continuation. Santa Monica, CA: Rand Corporation.

Hargreaves, A. (1990). Individualism and Individuality: Reinterpreting the Teacher Culture. (Paper presented at the Annual Meeting of the AERA, Boston).

Little, J. W. (1981). School Success and Staff Development in Urban Desegregated Schools: A Summary of Recently Completed Research. (Paper presented at the Annual Meeting of the AERA, Los Angeles).

Little, J. W. (1982). Norms of Collegiality and Experimentation: Workplace Conditions of School Success. American Educational Research Journal, 19(3), 325–340.

Little, J. W. (1990). The Persistence of Privacy: Autonomy and Initiative in Teachers' Professional Relations. Teachers College Record, 91(4), 509–536.

Nias, J., Southworth, G. & Yeomans, R. (1989). Staff Relationships in the Primary School: A Study of Organizational Cultures. London: Casell.

Rosenholtz, S. J. (1989). Teachers' Workplace. The Social Organization of Schools. New York: Longman.

Wilson, B. L. & Corbett, H. (1983). Organization and Change: The Effects of School Linkages on the Quantity of Implementation. Educational Administration Quarterly, 19, 85–104.

Zahorik, J. A. (1978). Teachers' Collegial Interaction: An Exploratory Study. The Elementary School Journal, 87(4), 385–396.

Zielinski, A. E. & Hoy, W. K. (1983). Isolation and Alienation in Elementary Schools. Educational Administration Quarterly, 19(2), 27–45.

THE TRAINING OF UNIVERSITY TEACHING STAFF: THE SPANISH CHALLENGE

Benedito Antoli, Ferrer Cerveró, & Ferreres Pavia, University of Barcelona, Barcelona, Spain

The teaching profession requires certain knowledge, skills and particular attitudes which not everybody has, and therefore these must be acquired in training preparatory to exercising the profession. Furthermore, one's exercise of the profession must be developed and improved.

But yet, in Spain, happens that, quite the contrary, one can gain access to the teaching profession in the University without any relevant qualification with regard to the specific functions required by professional development. One can become university teacher through selection procedures which have little or nothing to do with the real capacities that the profession demands. At best, there are attempts to guarantee that the candidate has a command of whatever knowledge is required by the discipline, which gives the impression that in order to teach one only has to know the subject. Thus, once faced with the responsibility of working with students in the classroom, the teacher is impelled to repeat certain forms of conduct and relationships that he has experienced as a student, that he has seen personified in the teachers he had, and in others who now share the experience of teaching with him.

It is true that there is not much idea of how to achieve the training education of university teaching staff. Amongst other reasons, there is a very complex framework of functions which develop the teaching /learning process in the classroom and out of it within an institution subject to serious contradiction and dilemmas: the university must prepare the student for the world of work, and yet, at the same time, has to make a critical reconstruction of culture; the university has to engage in research and simultaneously transmit the findings; the university must be in service of society and of the individuals going through it.

Spanish experience in the training of university teachers is not abundant. The conditions in which Spanish male and female university teachers have had to work, have not been favourable in Spain: a high ratio of students to teachers, few didactic resources, zero specific preparation, inadequate social prestige, low pay, little in the way of stimuli and not enough advantage is taken of the means of democratic participation.

At present, there is quite a lot of concern about this question of university pedagogy. In a great number of places in the world, we find diverse range of experiences. careful reading and interpretation of these facts should make us concerned not to commit errors that might put an end to these hopeful beginnings.

It is necessary the to avoid the contradictions that arise between what is required of the teachers and what is really done from the administrative point of view, in other words, the introduction of unsuitable training models, the rush to achieve in a limited period things which cannot be done in years, demands for obligatory change, lack of concern about providing high level conditions of work; and again, indiscriminate copying of models used in other countries (where they may have been successful); in sum, the putting into operation of some initiatives concerning which there has not been enough rigourous and intelligent reflection.

In accordance with these initial considerations, we would share in this paper the conclusions of a national report solicited from the general Section for Teaching Staff Training of the Spanish Ministry for Education and Science. The authors of this paper were members, among others, of the Working Commission that elaborated such national report on training university teaching staff.

The report " The training of university teaching staff" was developed during 1990-1991 and was finished in June 1991. It has been structured around six areas:

- Theoretical background concerning the functions and areas of concern of the university and of its teaching staff.
- X-ray view of the international and national situations with respect to teacher training.
- Professional development of university teaching staff.
- Teacher training strategies.
- Suggestions of an institutional nature.
- Follow-up and evaluation of proposals for improvement.

PROSPECTIVE MATHEMATICS TEACHERS' ATTITUDES TOWARD COMPUTER AND COMPUTER ASSISTED ACHIEVEMENT IN SOLVING MAX-MIN PROBLEMS IN CALCULUS

Zulbiye Toluk, & Yasar Ersoy, Middle East Technical University, Ankara, Turkey

Abstract

This paper deals with the determination of the prospective mathematics teachers (PMT) views on DERIVE, their attitudes and achievement. It is shown that there is a significant effect of DERIVE on the PMTs' attitudes toward computer and computer assisted instruction (C/CAI) and achievement in solving max-min problems in calculus.

Introduction

Computers continue to become an indispensable tool for technological society, and influence the school life in both developed and developing countries. In developing countries such as Turkey, there is however a little understanding of how to introduce computers effectively into classrooms because of various constraints, scare resources and a lack of experiences, Ersoy (1990). But it became a reality that the computer influences the way in which mathematics is taught or will be taught in near future, e.g. Howson and Wilson (1986). Recent developments, i.e. the increasing use of personal computers and mathematical software packages will have some consequences for both the contents of mathematics courses and the teaching/learning processes. The Computer Algebra Systems (CAS) e.g. DERIVE, MATHEMATICA performing symbolic manipulations can free individuals from doing algorithmic processes by hands so they can concentrate on higher level tasks such as solving applied mathematics problems. The purpose of the present study is to determine to some extent a group of the fourth grade students, i.e. the prospective mathematics teachers (PMT)' views on and the effectiveness of DERIVE introduced in an elective course, ScE 421 at the Dept of Science Education (DSE), METU in Ankara.

We have designed this investigation as a pre-pilot case study on the subject and continue to carry out various activities related to it. This pre-pilot case study indicates that DERIVE is effective in improving student's achievement in solving max-min problems, and changed the PMTs' attitude toward computers and

computer assisted instruction (C/CAI). The PMTs have enjoyed with working DERIVE, and had positive feeling toward the instruction on DERIVE.

Background Information

Development of Attitude Scales

Attitude is an important dependent variable which is affected by instructional method, and teachers' attitude toward C/CAI is a key factor in the successful implementation of computers in the classroom practice. The development of various attitude scales for students and teachers toward C/CAI and research on the modes of CAI have been of particular interest, e.g Kulik et al (1980), Roblyer et al (1988), Bulut and Ersoy (1992). For example, Kulik et al presented a meta analysis study which integrates the findings from 59 independent evaluations of computer based college teaching while Bulut and Ersoy developed an instrument to measure PMT attitudes toward C/CAI. PMTs should therefore come to understand what computer is and how it is used in teaching/learning process. These teachers may then play an important role in the computer assisted mathematics instruction (CAMI) at schools and probably have new tasks in changing the attitudes of students toward mathematics.

The Computing Environment

A systematical and widespread CAMI course for pre-service and in-service education of mathematics teachers has not been planned at teacher training colleges in Turkey eventhough programming courses such as FORTRAN and/or BASIC have been taught to PMTs in some universities, Ersoy and Toluk (1991). In 1988 a proposal was made for establishing computer laboratories for research on computer-based education (CBE), teaching some courses about programming and CBE to the prospective mathematics and/or science teachers at DSE, METU. This was the first CBE Project in the Turkish higher education and known as DSE-CALT Project, Ersoy (1991).

Method and Instruments

Problems and Hypotheses

The main problem is

"What is effect of DERIVE on the PMTs' achievement and attitudes toward C/CAI in ScE 421 course at METU?"

Because of the complexity the main problem will be more easily worked out if it is restated as subproblems (SP):

- SP1. What is the effect of the DERIVE on the PMTs' attitudes toward C/CAI?
- SP2. What is the effect of the DERIVE on the PMTs' achievement in solving max-min problems in calculus?

Subjects of the Study

The subjects are the fourth grade students, i.e. the PMTs, at DSE-METU trained for three weeks in ScE 421 course during the fall semester of 1991-92 academic year. Of the total 61 subjects the number of females is 35, and the number of males is 34.

Instruments and Procedures

Three instruments are used in the evaluation. They are Computer Attitude Scale (CAS), Mathematics Achievement Tests (MACH1 and MACH2), and an Opinionnaire. All these instruments were developed by Mathematics Education Research Group (MERG) at METU, Bulut and Ersoy (1992), Ubuz (1991). A pretest/posttest group design was used to gather the data in the present investigation. The subjects were distributed to three laboratory groups and had been trained for 12 hours, i.e. three weeks in ScE 421 course.

Analysis of data and results

Test of Hypotheses and Results

The hypotheses are tested for significance at a = 0.05 level, and the following results are obtained:

- H1: The result showed that the computer attitude mean score of the PMTs before the introduction of DERIVE into the course is significantly higher than the attitude mean score of the PMTs after using DERIVE.
- H2: Analysis show that MACH2 scores of the PMTs is significantly higher than the their MACH1 scores.

It was also found that there were no significant difference between the attitude mean scores of the male and female PMTs before and after the training period.

Views on DERIVE

The data collected by opinionnaire was analyzed by making a frequency count of responses to each question. Subjects, the PMTs, were asked to give their opinions about the uses of DERIVE in mathematics instruction in their professional life.

- About 90% of the PMTs found DERIVE useful in their professional life, and all of them enjoyed with working DERIVE.
- They pointed out that DERIVE should be supplemented with appropriate handout materials.
- More time and more exercise should be provided.

On the other hand, concerning the previous education of students for the course, 23% of the subjects determined their computer science background as poor. These students can be expected to have difficulties in ScE 421 course. Moreover, the responses to open-ended questions were also transcribed as follows. DERIVE:

- should be introduced in Math 151-152 courses, i.e Calculus;
- is very helpful in drawing graphs, and very useful in discovering the relationships between graph of a given function and its derivatives;
- can be used as a mathematical assistant in teaching calculus concepts; and
- leads students to be lazy in mathematical calculations.

Discussion and concluding remarks

From the analysis of data we found that DERIVE has a significant effect on the PMTs' attitude toward C/CAI and imply no significant difference between the male and female students' mathematics achievement (MACH1 and MACH2) scores. This result is consistent with the results obtained in the literature. Furthermore, it is interesting to notice that the attitude mean score of the PMTs before the introduction of DERIVE is higher than their attitude mean score after using

DERIVE. This result is inconsistent with the results given in the literature to some extent. More specifically, DERIVE:

- affects on PMTs' attitude toward C/CAI;
- changes student's ability to solve max-min problems; and
- aids students to discover the relationships between the graph of a function and its derivatives.

In order to identify the constraints and find out the reasons of inconsistency and changes in both attitudes and achievement various investigations on DERIVE should be designed.

References

Bulut, S., & Ersoy, Y. (1992, in print). An instrument for prospective mathematics teachers' attitudes toward computers and computer assisted instruction; METU J. Human Sciences , 10, Nr.2.

Ersoy, Y. (1990, May). Microcomputers into the educational system of developing countries: Proc of 7th Annual Congress of Computer Sciences., 199-204. Istanbul.

Ersoy, Y. (1991). The computer assisted learning/teaching project; Doga- Tr. J. of Physics, 15, 203-211.

Ersoy, Y., & Toluk, Z. (1986, in print). Pre-service education of secondary school mathematics teachers in Turkey; METU J. Human Sciences, 9, Nr.2.

Howson, G., & Wilson, B. (1986). School Mathematics in 1990s. Cambridge: Cambridge Uni Press.

Kulik, J. A., Kulik, C. C., & Cohen, P. A. (1980). Effectiveness of computer-based college teaching: A meta-analysis of findings. Review of Educational Research, 50, 11-16.

Roblyer, W. H. et al. (1988). Assessing the impact of computer-based instruction: A review of recent research. Computers in the Schools, 5, Nr 3/4.

Ubuz, B. (1991, unpublished). The Effect of Problem Solving Method with Handout Material on Achievement in Solving Calculus Word Problems; MSc Thesis, DSE-METU, Ankara,

EVALUATION OF A PRE-SERVICE TRAINING PROGRAMME BASED ON DIRECT INSTRUCTION

Simon Veenman, Yvonne Leenders, Pauline Meyer, & Mark Sanders, Department of Educational Sciences, University of Nijmegen, Nijmegen, The Netherlands

Based on the recognition that there is now a teaching effectiveness knowledge base of considerable range and depth that can be used for teacher education programmes and as a reaction to heightened concern about the quality of existing teacher education programmes, the Protestant Educational Advisory Centre and the Department of Educational Sciences of the University of Nijmegen, took the initiative for a pilot study to aid colleges of education in considering and assimilating relevant research on effective instruction in their attempts to improve.

Direct Instruction

Direct instruction as used in this study is a label for the constellation of effective teacher behaviours as synthesized by Rosenshine & Stevens (1986) from correlational and experimental studies. The core of this instructional model consists of six teaching steps or functions:

- daily review,
- presenting new material;
- guided practice;
- independent practice;
- weekly and monthly review;
- providing feedback and correctives.

This instructional model is most applicable to the teaching of performance skills or the teaching of well-structured learning-material (e.g. reading decoding skills, mathematical procedures, grammar, social studies facts). In this study this instructional model is called the direct instructional model for *explicit* skills. However, this model is less relevant for teaching in areas that are less well structured, that is, where all the steps cannot be specified and more than one answer is acceptable. Therefore, Rosenhine & Edmonds (1990) reviewed the existing literature on the teaching of less structured skills such as strategic reading, writing and mathematical reasoning. A major organizing concept for the teaching of implicit skills is that of scaffolding: instructional support which pupils receive form the teacher in order to help pupils bridge the gap between their current abilities and the goal. A scaffold or support is temporary and adjustable, and it is gradually withdrawn as the learners become more independent. Scaffolds may

include: modeling the process by the teacher, thinking aloud, provision of procedural facilitators, prompts, aids, guidance from the teacher, providing pupils with models to allow pupils to compare their work with that of an expert. Based on this review the instructional model for explicit skills was extended with the instructional elements that emerged from the implicit skill literature. This extension is called the instructional model for *implicit* skills. Both models are used in the course on effective instruction.

Research questions

The research questions that guided the study were:

- Do student teachers who participated in the course on direct instruction implement the desired teaching behaviours as presented in the two instructional models?
- Is there a positive effect on time-on-task rates in classes of student teachers who participated in the course?
- Do co-operating of supervising teachers of student teachers who participated in the course on direct instruction observe changes in desired teaching behaviours?
- How do student teachers and teacher educators value the new course on direct instruction?

Design

The study comprised three sub-studies:

- An observational study focusing on behaviour application by using trained observers;
- an observational study focusing on behaviour application by using ratings of supervising teachers;
- a questionnaire and interview study focusing on participants' reactions on the course of direct instruction.

The observational study focusing on behaviour application by using trained observers was set up as a field quasi-experimental pretest-posttest design with treatment (N = 27) and control (N = 14) groups of student teachers, and pupils associated with each student teacher. The observational study focusing on behaviour application by using supervising teachers was set up as an adaptation of the `pre, then, post' design (Mezoff, 1981).

Subjects

Participants in the study were student teachers from three colleges of education for primary teachers enroled in their second year of courses. College A was located in the north-west of the country, colleges B and C were located in the middle-east. Eight classes from these colleges were selected into the effective instruction course section (experimental or treatment groups). Student teachers from parallel classes, taught by the same teacher educators, did not follow the course on direct instruction (control groups). All student teachers took their education courses along with field experience. For logistical reasons a restricted number of student teachers in the treatment classes were randomly selected for the observational. After the student teachers completed the course on effective instruction the supervising teachers were asked to rate the performance of their student teachers. Of the mailed 150 supervising teachers, only 61 returned the scale (response rate 41%). Of the 205 questionnaires, used to get information on the student teachers'

perceptions of the content of the course, 130 were returned (response rate 63%). Interviews were conducted with three classes of student teachers and their teacher educators to collect information on the implementation of the course and suggestions for improvement.

Method and Instrumentation

Time-on-task

Pupils' time-on-task levels were collected by a time sampling procedure. Every 8 minutes during the lesson taught by the student teacher the observer stopped notetaking and recorded the number of pupils in the class who were engaged in academic activities or who were off-task. Each observation period lasted approximately 30 minutes and resulted in 3 time-on-task estimates. The inter-observer reliability for time-on-task checks revealed a coefficient of 0.94.

Direct Instruction Scale

After each observation, the Direct Instruction Scale (DIS) was used by the observers to assess student teachers behaviour on 19 variables (four-point scales that focused on the instructional skills). The DIS was based on the research of Rosenshine & Edmonds (1990), Veenman, Lem & Roelofs (1989). Inter-observer reliability checks for the separate instructional variables ranged from 0.72 to 1.00. After the student teachers completed the course on direct instruction, their supervising teachers were asked to complete the DIS. They were first asked to think back, and to rate the skills the student teachers demonstrated prior to the course (the `then' score). Second, they were asked to rate the same teaching behaviours after completion of the course (post scores).

Standardized lesson formats

To control for the influence of the lesson contents taught by the student teachers two types of lessons were drawn up: lessons for reading/language instruction and lessons for mathematics instruction. The mathematics lessons aimed at inviting the student teacher to the use of the direct instruction model for explicit skills. The reading/language lessons aimed at provoking the use of the direct instruction model for implicit skills. These two types of lessons (for grades 4, 5/6 and 7/8) were randomly distributed among the student teachers.

The instructional course

The course `Effective instruction' was implemented by teacher educators at three different colleges. Prior to the actual start of the course these teacher educators received a pre-training.

Data collection

Before the start of the course, each student teacher was observed during one reading/language or mathematics lesson (October - November 1991). After the course was provided each student teacher was again observed for one lesson (January - February 1992). Supervising teachers rated their student teachers after completion of the course. In the same period the evaluation questionnaires were submitted to the student teachers. The interviews with the four teacher educators

and three experimental classes (one class per college) also took place in this period.

Results

When examining the difference between the pre- and post-treatment data of the experimental group to determine if the treatment student teachers exhibited more of the desired instructional behaviours called for in the training programme, significant differences were found between the post-test and pre-test scores on the DIS. When inspecting the differences between the two subscales of the DIS, significant differences were found for the subscale Practice, but not for the subscale Presentation. The same results were found when examining the differences between the treatment and the control group (independent t-tests based on gain scores). The results of the study also indicated that the training programme had a significant effect on pupils' time-on-tasks rates. After the course, treatment group pupils exhibited significant increases in their time-on-task levels. The difference in gain scores between experimental and control group was also significant. Significant differences on the DIS scores were also found for the sample of supervising teachers: supervising teachers noted significant changes in desired teaching behaviours of student teachers who participated in the course on direct instruction. The instructional materials in the training programme were rated as particularly valuable by the student teachers. Almost all student teachers reported that the course was very helpful in learning to teach.

References

Mezoff, B. (1981). How to get accurate self-reports of training outcomes. Training and Development Journal, 35(9), 56-61.

Rosenshine, B., & Edmonds, J. (1990). New sources for improving instruction: The implicit skills study. Journal of Personnel Evaluation in Education, 4, 59-73.

Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock (Ed.), Handbook of Research on Teaching (pp. 376-391). New York: MacMillan.

Veenman, S., Lem, P., & Roelofs, E. (1989). Training teachers in mixed-age classrooms: Effects of a staff development programme. Educational Studies, 15(2), 165-180.

THE ETOBICOKE CLASSROOM ACTIVITIES PROJECT: THE MEANING OF EDUCATIONAL CHANGE

R. Volpe, R. Fleming, & A. Bennis, Canada

Summary

Inherent in the policy environment created by an emphasis on "learner centred teaching" is the need for participants to critically assess and problem solve in a continuous effort to move back and forth between policy and practice. What follows is a report of the accounts given by teachers on their reflections on what "learner centred" teaching means to them.

The purpose of this project was to develop an understanding of exemplary teaching practice and describe effective teaching strategies.

The focus of the project was on the ways effective programs contribute to the development of language and problem solving skills, a sense of self worth and efficacy, and the social abilities and consciousness of community that are demonstrated in good citizenship. Some of the specific areas of interest included ways of programming, reliable learning materials, the role of play, opportunities for activity, means of providing feedback to students, strategies for communicating to parents, and methods that work in dealing with special needs and their impact on the classroom.

Key Informant and Focus Group Interviews with teachers, administrators, students, and parents were supplemented with classroom observation. These in depth descriptions provided the basis for collaboration with teachers in the development and evaluation of new modes of student assessment.

What emerged from this study was a complex portrait of classroom effectiveness that was nested in personal experience and school – community processes. Subject matter competence, an interest in assisting student learning, sharing in discovery, a sense of appropriate outcomes for individual learners, the inclusion of opportunities for meaningful practice and recitation, a belief in the value of inquiry, the modelling of reflective practice, a flexible mix of traditional and non traditional teaching methods and an ability to share control of the learning processes with students all appear as patterns of what teachers and other stakeholders view as exemplary practice.

ON THE WAY TOWARDS ACTIVE LEARNING AND DEVELOPING WORKLIFE: DEVELOPMENT OF TEACHING IN FINNISH BUSINESSMEN'S COMMERCIAL COLLEGE

Antti Kauppi, & Seppo Peisa, Finnish Businessmen's Commercial College, Teacher Education Department, Helsinki, Finland

Problem Setting

The teacher education department of the Finnish Businessmen's Commercial College (FBCC) started in 1987 a quite broad and long development project, which aimed at developing the teaching and studying in the fields of commerce and administration. The project was based on an activity theory framework with an idea of learning as an activity. The project was strongly founded on for example Engestrîm's (1987) work on expansive learning.

Through the project some challenges for developing teaching and studying have become evident. The boundaries of school going have been much harder to break than was expected in the beginning. The students can't organize their studying towards real-life contexts and practical problems attached to them. Learning is therefore mainly directed towards school texts and examinations, not towards complex practical problems that would lead to applications in future work life. Students are still prisoners of the classroom that organizes their studying in 45 minute intervals in closed school systems. Meaningful and intensive learning activity is therefore hard to reach. The closed school system also is traditionally organized according to subject-matter boundaries. The subjects fill up the lessons and formal discussions, and the integration based on real objects in working life is hard to achieve.

The functioning of the school system seems to be based on time schedules and the definitions of teaching duties than on pedagogical premises. This reflects to division of labour, co-operation, and intensity of development work in the school. The institutionalized school context itself as a work environment in many ways prevents meaningful development of teaching and studying.

Consciously organized and supported development work has, however, pushed forward in FBCC some seeds for development that show possibilities for changing the teaching and studying strategies. These seeds can be found mainly in the teaching experiments that some of the teachers have conducted. This paper analyzes and examines teaching experiments conducted during the development

project to reveal seeds that somehow reframe the contradictions presented above and take teaching and studying forward.

Methodology

The development project in FBCC was organized around a methodology called developmental work research (Engestrîm 1987). The basic organization of the project has followed a sequence, where teachers' work has been researched, new teaching and studying strategies have been introduced, and practical teaching experiments have been implemented. Teachers themselves have participated in the development work.

The seeds of developing teaching practices have been most visible in some concrete teaching experiments. Therefore it seems to be necessary to analyze deeper the experiments and their theoretical and practical relevance. The analysis of the experiments leads to a more concrete and systematic elaboration of the seeds of development, and to generation of hypothesis for the zone of proximal development.

Analysis of Data

The teaching experiments have been planned and implemented by teams composed of supervising teachers and student teachers. The evaluation of the experiments has taken place through concrete material produced in the experiments including student teachers' reports and students' learning tasks, teachers' and students' interviews, and observation notes. The experiments have included an evaluation session, where the researcher, supervising teachers, student teachers, and students have discussed over the experiment. The concrete data includes recorded sessions and interviews, videotaped instruction, learning tasks, reports, and observation notes.

The data is very complex and diverse. The analysis has taken place through qualitative restructuring, where the researcher, supervising teachers, and student teachers have together elaborated the data. The essential tool in the data analysis has been modeling, which has been focused on the qualitative dimensions of the experiments. When analyzing complex and diverse data collected in real-life contexts the systemic viewpoint on concentrating on the whole activity systems has been of primary importance. The validation of the data and the results has taken place through the research objects' involvement in the data analysis.

Case: Closing of the Books in a Multinational Enterprise

To present an example of a teaching experiment, and its analysis, a course concerning the closing of the books in a multinational enterprise is elaborated further.

In Finland at the moment there are two different practices in closing of the books. On the one hand, the Finnish legal system requires a certain kind of closing of the books. On the other hand, there is an International Accountant Standards (IAS) - system that requires a different kind of closing of the books. This creates difficulties for the economic combinations between multinational enterprises and their subsidiaries as well as between co-operating partners in different countries. In practice the accountants have to move between these different practices in a fast

tempo, and make reports for both the Finnish interest groups and for the expert accountants of large multinational enterprises and international communities.
In the the experiment participated 1 supervising teacher, 6 student teachers, and 30 students. The learning task was so complex that even the supervising teacher and the student teachers were not well aware of the details of the problem area.
The experiment was organized in five phases:

- the student teachers and their students built up a study plan together;
- the students went into some multinational enterprises and interviewed their chiefs of accounting as well as accountants. The students collected material on e.g. the enterprise's ownership arrangements, reports of the closing of the books, copies of the official rules and regulations, and the enterprise's practices of accounting;
- the students built up reports of their inquiries, which were carefully discussed between student teachers and students;
- the students completed a learning task, where they had to change the closing of the books from the Finnish system into the IAS-system;
- the results were analyzed and discussed with an accounting expert from a multinational enterprise.

All together all teachers and students taking part in the experiment viewed the experiment as a very positive and effective learning experience. For example, one student argued that she had learned more during the experiment than during the rest of the two years that she had studied in the college.

Results

The main results highlight the need to expand the traditional notions of the object of learning, the subject matter, the student, and the way of approaching the reality.

- The starting point of the most positively experienced and effective teaching experiments has been the change of the object of teaching and learning from school text to complex and realistic economic situation.
- The subject matter in these experiments has been organized more on the basis of activity units, where the integration of different subjects is based on a real world object and its conceptual elaboration.
- The student was seen in these experiments more as an active and conscious subject, who is capable in analyzing studying and learning on a metacognitive level. This seemed to be at least partly due to the complex object of learning, which didn't give a possibility for simple examples and easy answers.
- The experiments made it necessary to approach the real-life economic situation through interventions and projects, where the student was actively involved in practical problem solving. This seemed to change the traditional and mechanistic school-going into an adventure and mutual inquiry in concrete and challenging environments.

References

Engeström, Y. (1987). Learning by expanding. Helsinki: Orienta-konsultit Oy.

INTERACTION STYLES AND COGNITIVE DEVELOPMENT

M. Lera, & A. Aguilera, Developmental and Educational Psychology, University of Seville, Spain

A great deal of research has related teacher behaviour and student performance (Brophy, 1983; Rosenshine, 1983; Mugny and Doise 1983; Perret–Clermont, 1984, among others). The present study analyzes the relationship between *teachers' cognitive activation patterns* and *student cognitive performance* (Mora, 1991; Aguilera, 1989). We understand the presence of certain teacher/student interaction patterns is a relevant variable in determining subjects' cognitive performance. The patterns are the following: process–oriented activity, reinforcement, self–concept stimulation, cognitive–conflict provoking, stimulation of alternative thinking, stimulation of explanatory thinking, stimulation of peer interaction, personal attention, slowing–down of class–dynamics, focusing, active–accomplice role, and increase in input. (Mora, 1987; Aguilera, 1989; Mora, 1991). We have observed that these teacher interaction patterns (Mora, 1985), displayed during a cognitive enrichment program, were accompanied by greater gains in scores on intelligence tests applied before and after treatment. We consider it necessary carry out research centred on the following aspects:

- the degree in which these cognitive activation patterns were present in teachers not participating in the program;
- if the presence of these patterns improves subjects' cognitive performance in standard classroom activities.

Method

Subjects

Twenty preschool classrooms (5 year olds) from the province of Seville (Spain), of different socio-economic status, were included. They were assigned to different level groups as a result of their scores on an academic performance test.
3 high–score classrooms, 14 medium score and 3 low–score classrooms were found.

Instruments

The "Tests de Preescolar de M Victoria de la Cruz" (1986) was used to assess cognitive development. This test measures verbal intelligence, mathematical concepts, auditive memory, motor–visual coordination, and perceptive aptitude and provides a total score for cognitive skills. A Spanish version of "The very hungry caterpillar" (Eric Clarke, 1986) was used for the experimental task. Sessions were videotaped.

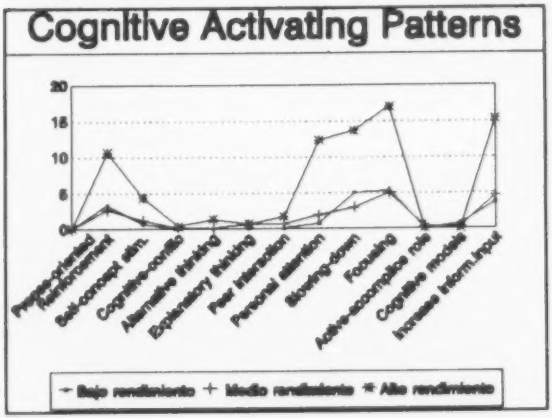
Procedure

Teachers were asked to carry out a task familiar to them: telling a story. "The very hungry caterpillar" was chosen. We gave them instructions to do it as they usually would in class, to take as long as they wished and that to accompany story–telling by whatever other activity they considered convenient, if they wished to do so later, the cognitive development test was administered (de la Cruz, 1986) to each class and, following classroom average scores, it was assigned to one of three groups: high, medium and low cognitive performance. Differences among groups were significant on Krushall–Wallis test ($p<.01$).

The analysis consisted of the identification, in one–minute intervals, of the presence of cognitive activation teacher–patterns that define a productive interaction pattern. The two observers reached a 93% reliability.

Results

Data analysis shows that all teachers only display certain cognitive activation patterns. All seem to have similar profiles but frequency is greater in groups with best results on cognitive performance test. "Reinforcement, personal attention, slowing–down of classroom dynamics, focusing and increase in input" are the most used (see chart 1).



Differences observed are significant (Mann–Whitney U test) on reinforcement, alternative thinking, centring, increase in input between medium and high level groups, and centring and increase in input between low and high level groups.

Discussion

Cognitive stimulating patterns do not spontaneously appear in teachers' standard practice. It therefore seems necessary to train them specifically in this type of interactions.

However, certain aspects can be observed in all teachers (reinforcement, personal attention, slowing–down of class–rhythm, focusing and increase in input). These patterns are related to class management/control activities, information–centred

activities and to demands specific to this age group. When significant differences are observed in these patterns children obtain different test results. Two alternative explanations are possible: on one hand, it could be that differences observed in child–teacher interactions are responsible for performance on the test, or, on the other, that student characteristics influence teachers' behaviour in such a way that with older children and less–able children the same teachers display less cognitive–enriching profiles. Further research, however, is needed to clarify this point.

References

Aguilera, A. (1989). Influencia de la interacción en la adquisición de procesos cognitivos. Influence of interaction on acquisition of cognitive processes]. Unpublished Graduate Dissertation. University of Seville.

Bennett, N., Jordan, J., Long, G., & Wade, B. (1976). Teaching styles and pupil progress. Cambridge MA: Harvard University Press.

Brophy, J.(1983). Classroom organization and management. Elementary School Journal, 83(4), 265–286.

Carle, E. (1990). La pequeña oruga glotona. [The very hungry caterpillar]. Madrid: Círculo de Lectores.

Cruz, M. V., de la (1988). Prueba de diagnóstico preescolar. [Preschool diagnosis test]. (2nd ed.). Madrid: TEA.

Mora, J. (1985). La estimulación de la inteligencia: Un programa de intervención por enriquecimiento cognitivo en sujetos deprivados socioculturales. [Stimulation of intelligence. An intervention program through cognitive enrichment in socioculturally deprived subjects]. Unpublished Graduate Dissertation. University of Seville.

Mora, J. (1988). Evaluación cualitativa del programa de Comprender y Transformar. [Qualitative evaluation of the program Understanding and Transforming]. Internal unpublished document of the Department of Developmental and Educational Psychology. University of Seville.

Mora, J. (1991). Evaluación del programa de enriquecimiento cognitivo Comprender y Transformar. [Evaluation of the cognitive enrichment program "Understanding and Transforming]. Unpublished Doctoral Dissertation. University of Seville.

Mugny, G., & Doise, W. (1983). La construcción social de la inteligencia. [Social construction of intelligence]. México: Trillas.

Perret–Clermont, A. N. (1984). La construcción de la inteligencia en la interacci:n social. Aprendiendo de los compañeros. [Construction of intelligence in social interaction. Learning from peers]. Madrid: Visor.

Rosenshine, B. (1983).Teaching functions in instructional programs. Elementary School Journal, 83(4), 335–351.

Soar, R. S., & Soar, R. M.(1979). Emotional climate and management. In P. Peterson & H. Walberg (Eds.), Research on teaching. Berkeley CA: McCutchan.

COGNITIVE ACTIVATING PATTERNS: AN INSTRUMENT FOR THE ASSESSMENT OF INTERACTIONS IN THE CLASSROOM

J. Mora, & A. Aguilera, Developmental and Educational Psychology, University of Seville, Spain

Reflection on cognitive enrichment programs and the consideration of data obtained with them shows that there are certain teaching skills that are responsible for the progress they produce in at least in the same proportion as content of the tasks (Mora, 1988a; Nickerson, 1986). For example, literature derived from research carried out around the "Instrumental Enrichment" program (Feuerstein et al., 1980; Feuerstein et al., 1988) is full of descriptions of mediational behaviours that are viewed as causes of cognitive change.

The study of the teaching skills that could enrich students' thinking opens an interesting field of research, but it requires new instruments that help identify the habilities providing most cognitive stimulation. This is the aim of our present investigation.

Method

Subjects

Three special education classes and two classes for culturally handicapped children were used. Teachers of all classes had administered the 'Understanding and Transforming' cognitive enrichment program (Mora, 1988b).

Procedure

Videotaped recordings of 26 program application sessions were analyzed, regularly distributed among the five groups. Bias due to contents of the unit being administered at each moment was eliminated by taking recordings at intervals of several months.

Assessment of program impact showed differential cognitive gains among groups. After an initial analysis, one of the authors developed a group of behaviour dimensions representative of stimulating teaching skills. This was submitted to the consideration of three independent groups of psychologists, who analyzed a sample of recorded sessions. With resulting data, a final system of twenty behaviour patterns was produced, describing teaching skills that facilitate cognitive enrichment.

These patterns were intended to form an open observation system, reflecting characteristics or aspects of meaning of behaviours, rather than a closed

conventional category system. To avoid methodological risks, contents of each aspect were written, together with examples of representative teaching behaviours. Observers were trained and final analysis was performed. Inter and intraobserver reliability was assessed.

Final observation was carried out registering presence/absence of 13 of these patterns per one–minute intervals. These were the only patterns apt for this type of observation. Global assessment scores for each of these and of the remaining seven patterns that required longer intervals for their observation were taken.

Results and Discussion hierarchical category system of stimulating teaching skills has finally been established as follows:

- process–oriented activity;
- reinforcement of cognitive behaviour of higher–than–base–level;
- stimulation of students self–concept;
- stimulation of cognitive conflict;
- stimulation of divergence and alternative thinking;
- stimulation of explanatory thinking;
- stimulation of peer interaction;
- personal attention;
- slowing–down of group–rhythm;
- group attention management;
- teacher adoption of an "active accomplice" role;
- teacher presentation of models of cognitive functioning by thinking "out aloud";
- increase in information input;
- fostering an active–participating climate;
- fostering an enjoyable climate;
- structuring the situation;
- flexibility in following proposed task;
- shaping, successive approaches to higher–order cognitive activity;
- presentation of problems to be solved in various sessions;
- fostering students' personal expression.

Relationship between teaching skills present in each group and students' cognitive gain measured with test criteria was analyzed. Data shows that the greatest amount of stimulating teaching skills is reached in the group with greatest gains, and that most of these skills follow the same proportions as student gains.

If proposed teaching skills are considered as a whole, it can be observed that the group of students with highest amount of stimulation is also the one that has reached greatest progress, and that stimulation in remaining groups is also proportional to gains.

The set of skills, as a whole, is therefore a good predictor of final achievement, and can be considered, from present exploratory perspective, to be responsible for it.

Data also shows how individually considered teaching skills are distributed among groups in the same fashion as cognitive gains in students.

Teacher process–oriented activity and presentation of cognitive models seem to be the decisive skills. Theory on cognitive stimulation has already pointed out the importance of these aspects (Nickerson et al., 1987).

There are two variables of little relevance, but that have distributed among groups in the same way as those considered critical: stimulation of alternative thinking and personal attention to thinking processes of each student.

Analysis of teaching skills carried out up to the present moment confirms our initial hypothesis that there are certain teacher attitudes that determine cognitive performance in students.

Results validate the analysis instrument used, since the group with most progress receives nearly two stimulating behaviours per one–minute interval, whereas lowest progress group receives one per two minutes. Proportion of stimulating behaviours in medium progress group is also intermediate.

References

Feuerstein, R., Rand, Y., Hoffman, M. B., & Miller, R. (1980). Instrumental Enrichment: An intervention program for cognitive modifiability. Baltimore: University Park Press.

Feuerstein, R., Rand, Y., & Rynder, J. E. (1988). Don't accept me as I am: Helping "retarded" people to excel. Nueva York: Plenum Press.

Mora J. (1988 a). Los programas de estimulación cognitiva [Cognitive Stimulation Programs]. Cuadernos de Pedagogía, 163, 60–62.

Mora J.(1988 b). El programa 'Comprender y Transformar'. [The 'Understanding and Transforming' program] In A. Fierro Psicología clínica. [Clinical Psychology]. Madrid: Pirmide.

Nickerson, R. .(1986). Project Intelligence: An account and some reflections. In M. S. Schwebel & C. A. Maher, Facilitating Cognitive Development: International perspectives, programs and practices. London: The Haworth Press.

Nickerson, R. S., Perkins, D. N., & Smith, E. E.(1985/1987). Enseñar a pensar: Aspectos de la aptitud intelectual. [The Teaching of Thinking]. Barcelona: Paidós/MEC.

DYNAMIC COLLABORATIVE APPROACHES TO STAFF DEVELOPMENT REVIEW AND NEEDS IDENTIFICATION AT SCHOOL AND EDUCATION AUTHORITY LEVEL IN SCOTLAND

Arthur Naylor, & James P. O'Brien, ST Andrew's College, Beardsen, Glasgow, United Kingdom

Background and remit of the research project

In November 1991 St Andrew's College was approached by the Educational Development Service of Lanark Division, Strathclyde Region to determine whether the college could devise a mechanism to assist in the identification of staff development needs in secondary schools across the Division for session 1992-3. Lanark Division is one of six divisions of Strathclyde Region which is the largest education authority in Great Britain. There are 38 secondary schools in Lanark

Project Aims

The aims of the research project were:

- to review inservice needs in secondary schools currently being addressed;
- to determine the extent to which such needs may still require to be addressed;
- to identify new inservice needs for the next session;
- to be specific about topics and themes which require to be overtaken in relation to these needs.

Timescales

The project was divided into four phases:

- Involvement of divisional education advisers, headteachers and newly appointed school staff development co-ordinators. At each stage of involvement there was an opportunity to influence the design of the process and research instruments to be used. Nominal Group Technique methodologies and Rating Scale Questionnaires produced by Jim O'Brien and Arthur Naylor were shared and discussed at meetings with advisers, headteachers and staff development Co-ordinators.
- In January and February 1992 a team of eight college staff worked with, on average, five schools each to explain the processes and the rating scales.
- Questionnaire returns from schools involving responses from each subject area and from school management were obtained from 35 out of 38 schools - a response rate of 92.2%.

- Analysis, illustration and reporting of the data to the Division took place in March 1992. A written report has been provided for the Division and individual reporting back is being arranged for each individual participating school.

Outcomes

The major needs identified by subject principal teachers were:

- 5-14 Curriculum and Assessment;
- Revised Higher grade;
- Information technology;
- Middle Management Training;
- Classroom methodology;
- Development Planning;
- Primary/Secondary liaison;
- Student-centred Learning;
- Learning support.

Particular value was placed on school-based delivery of staff development working where appropriate with an external consultant rather than on centre-based courses.

There were significant variations on the needs expressed by principal teachers of different subjects and variations between the highest identified needs as perceived by classroom teachers and education authority advisers.

CREATIVE CLIMATE IN TEACHERS' GROUP: BARRIERS AND CHANGES OF INNOVATIONS

Daniela Kusá , Institute of Experimental Psychology, Bratislava, Czechoslovakia

The ability to produce unusual, creative solutions is traditionally connected with independence, nonconformity, and risk - taking in creative persons. But, research on a creative climate in work groups has indicated "face" of creativity in an individual is merely one of the conditions that are favourable for creativity in a group (Ekvall, G., & Andersson, Y., 1986; Isaksen, S.G., & Kaufman, G., 1991; Nystrm, H., 1991).

In creative climate, dimensions are pursued involving a measure of individual independence (freedom, risk - taking), but dimensions expressing a measure of social dependence are also involved (support, trust). In other words, a group forms the field of forces that operate for/against the benefit of individual creativity.

If we consider the barriers of creativity, there are not only the barriers with a cognitive content such as the lack of stimuli, discussions, dynamism in production of different solutions, but also the barriers with a social content such as lack of trust, support, amount of conflicts, etc..

What is the importance of particular dimensions in teachers' groups, where results of work are delayed? How to improve the teachers creativity in his/her immediate work? What are the specific conditions that need special attention?

Method

The sample was formed from 40 teachers of the primary school. There were used the methods:

- CCQ - Creative Climate Questionnaire /Isaksen's modification of the Ekvall method/ involving 10 scales: challenge, freedom, support, dynamism, playfulness, debates, trust, conflicts, risk - taking, and work pace;
- Category Width Questionnaire /Pettigrew/: category width as a supposed measure of cognitive risking/carefulness;
- Choice Dilemma Questionnaire /Kogan, Wallach,/: decision making in 12 model situations providing the risky /careful alternatives;
- Stroop's Interference Test: the resistance to perceptional interference was measured as an index of work pressure at the end of week.

Results

CCQ dimensions - evaluation and inner relationships:

- Debates and Challenges were shown to be the highest evaluated dimensions. However, regarding the possible scale values in this method, these two dimensions were found to be within average range of evaluation;
- Work pace, Conflicts, and Trust were shown to be the lowest evaluated dimensions. Thus, high work pace, prestige, and luck of trust seemed to be the barriers of a high importance;
- as to inner relationships, the highest correlations were found among dimensions Support, Dynamism, Playfulness, and Trust (.635 to 737). Dimensions Freedom and Debates were shown in the lowest significant relationship (.397).

Risk - taking indicators:

- regarding Risk - taking in CCQ, the only significant relationship to dimension Challenge was found (.628). Other relevant relationships were not shown (to Freedom, Dynamism, etc);
- category width and decision making in risky/careful alternatives were found in a significant relationship (.403);
- no relationship between category width/decision making and CCQ dimensions was shown;

Work pace in Stroop Interference Test and CCQ:

- in interference test, no significant differences were found between the teachers' performance and the norms. Thus, speed and accuracy in interference load was on a mean level;
- regarding Work pace in CCQ, the relationships to Debates (.441), and to Challenge (.397) were found. As the only dimensions in CCQ, Work pace and Debates were in no relationship to an individual opinion about an acceptance of his/her creativity in a work group.

Discussion

The research on a creative climate in work groups is connected with looking for the ways to increase ideas and innovations, to improve direct results of work. But, a creative climate could also be considered an indicator of satisfaction in his/her professional life. Of course, the direct relationship between a creative production and satisfaction with work cannot be expected. On the other hand,if the global evaluation of a creative climate is on a low level, dissatisfaction such as stress can operate against creativity.

In the research presented here, work pace was evaluated most negatively - it approached zero value. However, perceptional performance has shown that teachers' load was not extremely high - it was not on the level perceived by teachers.

In a global profile of a creative climate, no dimension was found that could counterbalance negative evaluation of work pace. Beside dimension Debates, the other dimensions were evaluated below average of possible scale value, whereby the lack of trust and high level of prestige could contribute to the perception of work pace as being extremely high.

"Both relaxed and stressful conditions can produce increased performance. Some arousal is necessary for creativity, but too much stress is detrimental. In the same way, conditions may become too playful, so that no work is accomplished ..." (Torrance, E.P., 1988, p. 65)

In the research on highly innovative groups, work pace evaluated at different levels, its placement was not exceptional (Ekvall, G., & Anderson, Y., 1986). In groups with different level of innovations, work pace was also not found to be important (Nystrm, H., 1991).

Teachers are not commonly regarded as a group with a high level of innovations. However, results have indicated dissatisfaction of teachers with creative climate, which probably influenced the evaluation of work load in a negative way. At the same time, competitiveness consisting in low level of trust and an increased focus on prestige can be the source of the load as well as a creativity barrier in teachers.

References

Ekvall, G., & Anderson, Y. (1986). Working climate in creativity: A study of an innovative newspaper office. J. of Creative Behaviour, 20, pp. 215-225.

Isaksen, S. G., & Kaufman, G. (1991). Adaptors and innovators different perceptions of the psychological climate for creativity. In Creativity and Innovation: Learning from Practice. Innovation Consulting Group TNO.

Nystrm, H. (1991). Organizational Climate and Innovative Capacity: A Comparative study of Three Companies. In 2nd European Congress of Psychology, Abstracts Vol. I., Budapest.

Torrance, E. P. (1988). The nature of creativity as manifest in its testing. In R. J. Sternberg (Ed.), The Nature of Creativity. (pp. 43-75). Cambridge University Press.

TRAINING FOR COMPETENCE IN TEACHING – ASSURING COMPETENT PERFORMANCE IN PRACTICE

F. Kleviet, & H.H. Tillema, Centre for the Study of Education and Instruction, University of Leiden, Leiden, The Netherlands

Recent theoretical and research contributions in the field of cognition and instruction have evolved in a knowledge base that may have clear implications for rethinking the training and preparation of teachers. Advances in studies on for instance knowledge acquisition and transfer (Salomon & Globerson, 1989), meta–strategies (Chi & Bassok, 1989) or intervention studies, bringing about specific learning results (Resnick, 1989), can provide new impetus to the design of teacher–training and help teachers to acquire new insights for classroom practice. New conceptions on instruction as is the case in the interest for situated learning (Collins et al. 1989) can also be of consequence to the field of teacher training. It is the focus of this symposium to highlight developments in cognitive theories of instruction that have possible impact in teacher training. It can be asserted that these developments have a high potential for a better foundation of the way teachers are trained for their profession. As such the different contributions to this symposium all deal with this focus but from different perspectives:

- the design of training programs for teachers in order for teachers to learn using certain teaching approaches (Bossink);
- the transfer of training in order to see whether teacher's learning can be influenced in such a way that utilization of concepts is warranted in practice (Börger);
- the assessment of teaching competence in order to relate knowledge and performance (Verloop).

The introductory contribution itself deals with three problem areas:

A cognitive approach to training of teachers

Several studies have been conducted, most of them explicitly designed as intervention studies to test cognitive theories in a constructive way. For instance, researchers as Carpenter, Schoenfeld, Bereiter, Lampert, (see Collins, 1989) all have used treatment procedures in classrooms to teach subject matter to pupils. The subject matter was explicitly taught according to principles derived or extracted from a cognitive theory. These studies work in two ways: on the one hand they give opportunity to test a cognitive hypothesis; on the other hand they give way to an instructional design solution for teaching the subject matter according to the theory. In this respect these studies pertain to the required competence and instructional behaviour of teachers and indirectly to the competence of teacher to work with these cognitive principles in the classroom.

949

Some of these studies make use of expert teachers, others use specific training to acquaint teachers with certain teaching approaches. These cognitively guidedforms of training seem highly promising in that they convey new approaches to teaching in order for teachers to work with this knowledge in applied settings (situated as it were). Important principles by which these new approaches to training operate are: clear conceptual understanding of the instructional behaviour, peer review and communication about teaching approaches, working out examples and designing specific lesson arrangements. Some of the approaches are:

- conceptual change approach to teaching in which teacher perceptions and beliefs are taken as starting point for introducing new concepts or strategies in teaching;
- cognitively guided instruction in which research based teaching strategies are introduced, discussed and practised by teachers;
- teaching study groups in which self study and collaborative designing learning environments help teacher as designers of instructional environments;
- conceptual training in which acquisition of teaching competencies is conducted in a highly explicit manner.

Assessment for competent teaching

Training implies that somehow performance levels can be established. After a program has been delivered change in mastery must be established in order to prove effects. Teaching is however a complex activity not easily screened by simple means of testing, especially when it is concerned with change or gain in knowledge–levels only; adequate measurement of knowledge domains and competencies is needed with respect to the performance level and mastery of competencies in real settings with real classroom tasks. This calls for more direct forms of measurement that authentically grasp professional expertise.
New forms of measurement have been developed (Shulman, 1990, Bird, 1990) such as:

- portfolios in which specimen of classroom actions are collected;
- assessment methods (such as devising diagnostic strategies for pupils with learning difficulties, cooperative planning assignments, decision–making with vignettes and simulated of teaching tasks) have been developed.

The design of these methods but remain problematic with respect to reliability and generalizability (Haertel,1990). Criteria need to be developed by which these forms of measurement can be evaluated such as content quality, complexity, meaningfulness but also replicatability and costs.

Transferring knowledge to classroom contexts

One of the central problems in the acquisition of content–learner– and pedagogical knowledge from educational theory and instructional research (Carpenter et al. 1991) is that teachers find it difficult to transfer this knowledge to their classroom; put otherwise, this educational knowledge is reluctantly integrated with the practical or craft knowledge of the teacher (Kieviet, Boei and Corporaal, 1990). Transfer for practice can be regarded as building on sufficient domain specific knowledge that can become more easily accessed in practice situations (i.e. a knowledge organization problem– Borger and Tillema, 1991). Another interpretation is regarding transfer as a problem of lacking in situational cues that

call for utilizing knowledge in the right context. Both interpretations of the transfer problem lead to wholly different solutions. Studies on transfer give reason to contextualize knowledge acquisition in more rich learning environments in order to reach greater similarity between the context of acquisition and of utilization. There are also indications that problem solving strategies and strategies as monitoring and self explanations can integrate knowledge with subsequent action
Studies in the different areas reviewed, give new impetus to the design and arrangement of more realistic environments in which teachers can learn the competencies needed for their profession. It seems that there are now promising leads available to revitalize the body of knowledge on teaching strategies and learning to teach.

References

Bird, T. (1988). The school teacher's portfolio: Assessment project. Stanford University.

Börger, H., & Tillema H. H. (1992). Transferring knowledge to classroom teaching: Supplying educational theory to student teachers. In C. Day, J. Calderhead & P. Denicolo, Research on teacher thinking, towards understanding professional development.

Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study, American Educational Research Journal, 26(4), 499–531.

Chi, M. T., & Bassok, M. (1984). Learning from examples via self explanations, in L.B. Resnick ed. Knowing learning and instruction, New Jersey: Lawrence Erlbaum

Collins, A., Brown, J.S. Newman S. (1989) Cognitive apprenticeship, teaching the crafts of reading, writing and mathematics. In L. S. Resnick, Knowing learning and instruction. New Jersey: Lawrence Erlbaum.

Haertel, E. H. (1990). Performance tests, simulations and other methods, In J. Millman & L. Darling Hammond, The new handbook of teacher evaluation. (pp. 278-295). London: Sage London.

Kieviet, F., Boei, F., & Corporaal., A. (1992, in press). Reflections upon 5 years of repertory grid research into declarative schemata of prospective teachers, Journal of Structural Learning.

Prawat, R. S. (1989). Promoting access to knowledge strategy and disposition in students: a research synthesis, Review of educational research, 59(1), 1–41.

Shulman, L., Haertel, E., & Bird, T. (1988). Toward alternative assessments of teaching: Teacher assessment project. School of education, Stanford University.

Shulman, L. S. (1986). Those who understand, knowledge growth in teaching. Educational researcher, 15(2), 4–14.

Smith, D. C., & Neale, D. C. (1989). The construction of subject matter knowledge in primary science teaching, Teaching & Teacher Education, 5(1), 1–20.

CONSTRUCTION OF A CONCEPTUAL TEACHER TRAINING METHOD BASED ON CONCEPTUAL EXCHANGE

C. Bossink, Centre for the Study of Education and Instruction, University of Leiden, Leiden, The Netherlands

Introduction

Acquisition and development of student teachers' professional competence depend on the available knowledge; the pre–existing (declarative) knowledge base (so–called "prior knowledge" or "preconceptions") acts as a framework within which new learning takes place (Shuell, 1986). Studies on learning to teach in different subject–matter domains (Carter, 1990) reveal that student teachers possess substantial pedagogical and subject–matter knowledge. Nevertheless, this pre–existing knowledge may be inconsistent with the essential concepts in a discipline or, for example, from notions about good teaching in a certain domain. Preconceptions may be overgeneralisations, alternative conceptions or even misconceptions. When such a *conceptual discrepancy* exists, student teachers often maintain their original conceptions about teaching and about content when presented with new information, for example in teacher education courses. Thus the existing knowledge base may interfere with new learning and thereby undermine the influence of the teacher education institute.

This study aims to investigate the role of instruction in influencing the pre–existing knowledge of student teachers (Corporaal & Boei, 1990).

Traditional training methods often have great difficulty to overcome student teachers' resistance to changing or restructuring their existing conceptions. In these methods little attention is paid to the, nevertheless, well–known educational principle "taking account of prior knowledge..." (Fiske & Taylor, 1984). Even if these methods would, in the case of alternative conceptions or misconceptions a – more or less – radical change in pre–existing knowledge might be necessary to prevent interference with new learning. This requires active intervention directed to bringing about this change.

Using a quasi–experimental design, the effect of two conceptual instruction variants – direct instruction and conceptual exchange – on the declarative knowledge base of student teachers will be tested. These instruction methods differ in the extent to which the presentation of new concepts is adjusted to students' preconceptions.

Direct Instruction

The method of direct instruction (also called "effective" or "explicit" instruction) represents more traditional training methods. This method is based on a synthesis of several teacher behaviours that were found to be effective according to process–product research (Rosenshine & Stevens, 1986). Major components of this instruction method include teaching in small steps with student practice after each step, guiding students during initial practice, and providing all students with a high level of successful practice. Instruction is guided by the subject matter the teacher wants to convey to his students. "Taking account of prior knowledge..." is limited to a daily review and check of previous day's homework.

952

Conceptual exchange

By conceptual exchange we mean an exchange between the pre–existing (declarative) knowledge of students and the knowledge offered by the teacher education institute (Corporaal, 1988). In an instruction method based on conceptual exchange, (1) students' preconceptions are elicited and assessed (by asking students for explanations and predictions about phenomena), (2) preconceptions are contradicted by confrontation with discrepant information, and (3) the students' constructions of new concepts are facilitated (Eylon & Linn, 1989; Smith & Neale, 1989).

Conceptual exchange training of teachers

Teachers from primary teacher education institutes will be trained in teaching educational theory (conceptual exchange and direct instruction) to their own students according to both instruction methods. The operationalization of conceptual exchange teaching will be based on that of Neale, Smith, and Johnson (1990), who focused on teachers' pedagogical content knowledge concerning science. Their model includes a relatively specific blueprint for the structure of classroom lessons and the appropriate roles for students and teachers. Much of their success in relating teachers' performances on questionnaires or experimental tasks to the real classroom environment may be attributable to what the researchers chose to measure:

- pedagogical beliefs specific not only to a content field but also to certain topics within that field;
- pervasive pedagogical orientations that affect every aspect of a teacher's classroom practice;
- a very specific instructional model;
- relatively long–term change in beliefs (Kagan, 1990).

From the beginning of this training period, teachers are actively involved in designing manuals (for both training methods) to be used by student teachers in their own practice class–rooms. Furthermore, teachers read and discuss research on alternative conceptions and on teaching strategies. Readings are chosen to provide evidence that contradicts teachers' views. Teachers also test their own subject–matter knowledge in activities that revealed their alternative conceptions and provided opportunities for them to construct more adequate conceptions. Before, during and after this training program, data will be obtained from multiple sources, including structured interviews, teachers' self–evaluations, and transcripts of actual classroom lessons. Interview data will be used to infer general profiles of each teachers' orientation to teaching educational theory and to assess teachers' knowledge and beliefs about subject–matter knowledge and pedagogical content knowledge.

This training program will be evaluated and presented in more detail at the conference.

References

Carter, K. (1990). Teachers' knowledge and learning to teach. In W. R. Houston (Ed.), Handbook of research in teacher education (pp. 291–310). New York: Macmillan publishing company.

Corporaal, A. H. (1988). Bouwstenen voor een opleidingsdidactiek.: Theorie en onderzoek met betrekking tot cognities van aanstaande onderwijsgevenden. De Lier, The Netherlands: Academisch Boeken Centrum.
Corporaal, A. H., & Boei, F. (1990). Instructie en de ontwikkeling van declaratieve kennis bij aanstaande onderwijsgevenden. In Th. Wubbels, Th. C. M. Bergen & F. A. J. Korthagen (Eds.), VELON jaarboek. recent onderzoek over onderwijs en opleiding (pp. 13–26). Utrecht/Groningen, The Netherlands: W.C.C./VELON.
Eylon, B., & Linn, M. C. (1988). Learning and instruction: An examination of four research perspectives in science education. Review of Educational Research, 58(3), 251–301.
Fiske, S. T., & Taylor, S. E. (1984). Social Cognition. New York: Random House.
Kagan, D. M. (1990). Ways of evaluating teacher cognition: Inferences concerning the goldilocks principle. Review of Educational Research, 60(3), 419–469.
Neale, D. C., Smith, D. C., & Johnson, V. G. (1990). Implementing conceptual change teaching in primary science. The elementary school journal, 91(2), 109–131.
Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M. C. Wittrock, Handbook of research on teaching (pp. 376–391). New York: Macmillan publishing company.
Shuell, T. J. (1986). Cognitive conceptions of learning. Review of Educational Research, 56(4), 411–437.
Smith, D. C., & Neale, D. C. (1989). The construction of subject matter knowledge in primary science teaching. Teacher & Teacher Education, 5, 1–20.

TRANSFERRING KNOWLEDGE TO CLASSROOM PRACTICE: PROVIDING CONTEXTUALISED KNOWLEDGE TO STUDENT–TEACHERS

M.A. Börger, Centre for the Study of Education and Instruction, University of Leiden, Leiden, The Netherlands

Relevance

In teacher– and inservice education the criticism is frequently encountered that educational programs do not provide knowledge that can be readily put into practice. This is confirmed by the finding that student–teachers who do have the required knowledge do not use this knowledge in practice. This lack of usability or practice potential can be regarded as a problem of transfer of knowledge to relevant practice situations. Therefore it is a problem of learning and thus alterable through instructional arrangement of (teacher education) programs. This kind of knowledge transfer to adequate classroom application has however been the subject of a surprisingly small amount of research within teacher educational studies.

Knowledge utilisation

Recently utilisation of the professional knowledge owned by teachers has gained some attention. Two didactical approaches can be mentioned that give guide-lines to the instruction of theory, promoting the use of it.

The *embedding approach* focuses on the accessibility of knowledge by linking rules to knowledge, and adequately structuring knowledge. Strategies are conceived as handles of knowledge domains. Through these handles domain–specific knowledge can be retrieved and used (Perkins & Salomon, 1989). New information should be structured and related to existing knowledge, leading to interrelated knowledge structures. Adequately structured knowledge is less inert, leading to a more flexible application of knowledge in practice situations (Voss, 1987).

The *immersion approach* argues that theoretical knowledge should be immersed into the practical context. Learning in their view is guided by ideas that emerge from the context in which learning takes place (Prawat, 1991). New information should therefore not be explicitly taught to students, but students should be stimulated to generate rules and ideas out of the learning context. The context of learning should thus be as "real" and as "rich" as possible (see for example The Cognition and Technology group at Vanderbilt, 1990). In this way experience based and domain specific strategies are learned. The learned knowledge will therefore be more meaningful and more likely to be cued by situational events within the applicational context.

The interaction of strategies and knowledge–base

Knowledge acquisition theories traditionally concentrated on the aspects of knowledge content and organisation, emphasising the importance of similarity between the learned task and learning environment, and the later task performance and applicational context. Later the importance of knowledge independent processes like the use of metacognitive strategies also became clear. These strategies were expected to enhance transfer because of their generality. Nowadays both aspects are believed to influence the process of transfer: "structures of knowledge and cognitive processes interact" (Garner, 1990). The above mentioned didactical approaches have different views about the relation between these two aspects. As earlier mentioned, the embedding approach emphasises the importance of the knowledge organisation. General and more domain specific strategies have to be linked with this knowledge base in order to function as operators. In the following study this is operationalised by directly linking rules and examples to –small amounts of– the learned knowledge. The immersion approach on the other hand tells us that learning by using metacognitive strategies will lead to generating domain specific strategies and knowledge. This will be operationalised using the method of *self–explanation* (see for example Chi & Bassok, 1984); Examples of the theory in practice are given, and the students are stimulated to give as much explanation for the didactical actions as possible. Using metacognitive and monitoring strategies applicational strategies are directly abstracted out of the classroom context.

Research proposal

A research study will be undertaken to compare the two basic instructional methods, underlying the above made distinction. Subjects are students in their second year of a four–year training to become a primary teacher. The content of the educational theory they received is the Advance Organiser theory of Ausubel. Twenty students are instructed using a highly structured rule–example method and twenty students receive an instruction based on self explanation of examples. Both groups receive the same amount of instruction and exercises (total duration: three hours). Data are gathered in the form of questions (strategic and more factual) about the theory–content and in the form of a knowledge map representing the theory learned.

The students are then ordered to give a lesson according to the learned theory. Their lesson–plans are analyzed on the number of theory related elements. A small sample from each group is taken whose lessons are taped on video. Stimulated recall immediately follow these lessons. The lessons and reports are also analyzed on the number of theory–related elements.

The training material for this study will be discussed.

References

Chi, M. T. H., & Bassok, M. (1984). Learning from examples via self–explanations. In L. B. Resnick (ed.), Knowing, learning and Instruction. New Jersey: Lawrence Erlbaum.

Perkins, D. N., & Salomon, G. (1989). Are cognitive skills context–bound? Educational Researcher, 18, pp 16–26.

Prawat, R. S. (1991). The value of ideas: The immersion approach to the development of thinking. Educational Researcher, 20, pp. 3–11.

The Cognition and Technology group at Vanderbilt (1990). Anchored instruction and its relationship to situated cognition. Educational Researcher, 19, pp. 2–11.

Voss, J. F. (1987). Learning and transfer in subject–matter learning: A problem solving model. International Journal of Educational Research, 11, pp. 607–622.

ASSESSMENT OF TEACHER COMPETENCE

Nico Verloop, Teacher Education Department, University of Leiden, Leiden, The Netherlands

Introduction

During the last ten years there has been a growing interest in teacher evaluation, although there has been a considerable variation between countries in this respect. Without doubt the increased awareness of the teacher's crucial role in education has stimulated this interest. The focus of attention has gradually shifted from curricular and innovative matters to the person of the teacher as a factor of decisive importance. Doubts about the quality of the teaching force in general and a growing emphasis on accountability further stimulated discussions about teacher assessment.

A second, be it less explicit, impetus to the growing interest in teacher evaluation, is the increased awareness of the professional character of teaching.

Evaluation practices

Up to now teacher evaluation, if practiced at all, was done mainly by either written examinations or by observation of lessons. Criticisms with respect to evaluations focus at the validity of both procedure and criteria, even if these criteria are based on research evidence (for a related discussion concerning criteria form 'direct teaching' –research see Haertel, 1991).

Recently a number of promising developments with respect to teacher evaluation in the future have emerged. The related evaluation procedures can be situated between the one extreme, paper and pencil evaluation, and the other one, unstructured observation in the classroom.

One common characteristic of these procedures is the striving for a more valid measuring of the multifaceted teacher task. Many of these procedures are inspired by assessment procedures from other occupational groups, e.g. management assessment centres.

A second common characteristic is a growing tendency to bring these assessment procedures under the control of the teachers as professionals. Instead of an 'external' bureaucratic activity, assessment is seen as a responsibility of the teaching profession.

Indicative of this development is the founding in the USA of the National Board for Professional Teaching Standards, whose mission it is to develop a voluntary certification system for experienced teachers. The majority of the Board consists of elementary and secondary schoolteachers. By way of illustration some of these evaluation instruments, developed by Board order, are mentioned here (cf. Bird, 1990; Haertel 1990):

- Instructional vignettes. The teacher is confronted with a number of (life or video) problem situations in e.g. elementary arithmetic, and is asked how exactly he or she would respond;
- Topic sequencing. The teacher is confronted with a number of related topics from a particular domain and is asked all kinds of questions regarding its selection, sequencing et cetera, in a series of lessons;
- The teacher's portfolio. This concerns a teacher's personal file, containing all kinds of information (observational notes, pupil feedback, class handouts, audiotapes of lessons, et cetera) about the teacher's behaviour. This information could provide additional data for discussing and evaluating the teacher's functioning.

Problems and prospects

The most important problem with respect to all these measures concerns their validity. The fact that this type of assessment, like all performance assessments (cf. Mehrens, 1992), has face validity, cannot conceal the fact that there are some real validity problems, the most salient one being the lack of an unambiguous criterion. Some seem to suggest that this problem can be solved in a straightforward manner by starting from a job analysis of the teacher's job (e.g. Wilson, 1989). Others (like Haertel, 1991) contend that we should not start form the 'average' or typical classroom situation, but instead take the outstanding, exemplary teacher as the criterion. In view of the fact that these instruments pretend to assess

professional behaviour of experienced teachers, the choice for the latter option seems obvious to me.

There might be a crucial role for cognitive psychology here. During the last decade the amount of research on behaviours and cognitions of experts was considerable, not only pertaining to expertise in general (Glaser, 1986), but also specifically pertaining to expertise in teaching (Sabers, Cushing & Berliner, 1991). The next step could be incorporating these insights into an evalution system for experienced teachers. No doubt this will, depending on the specific evaluation tasks at hand, require additional research efforts into details of the experts' thought processes.

Another crucial problem concerns the question whether this type of evaluation will be acceptable to the teachers. At this moment there are some research results with respect to teachers' opinions about various aspects of teacher evaluations systems. It appears that in general the purpose of the evaluation determines the teachers' opinions about the whole system. Research from both the USA (Peterson & Cameaux, 1989) and de UK (cf. Powney, 1991) revealed that teachers saw the ideal purpose of teacher evaluation as one of promoting professional development and encouraging teachers to reflect and think about their practice. In other words, as long as the appraisal system is perceived as being formative in nature, teachers generally are willing to cooperate, while viewing the appraisal system as a means for bureaucratic control might lead to teacher resistance and apathy. In my opinion, given the many unresolved problems in the field of validity and reliability at this moment, it seems wise to confine the function of the newly developed evaluation procedures to a formative one. In this way teacher evaluation might also serve as a means for putting an end to the isolation in which many teachers are working, by providing a format for discussing and evaluating their classroom activities. In view of the fact that also the teachers themselves definitely see the point of reflection on their own functioning, there seems to be another task for cognitive psychology, (especially with respect to the development of competence in self–regulation and reflection). It should be investigated whether the insights from this line of research might be incorporated into an evaluation system for teachers aimed at professional development.

References

Bird, T. (1990). The schoolteacher's portfolio: An essay on possibilities. In J. Millman & L. Darling–Hammond (Eds.), The new handbook of teacher evaluation. London: Sage.

Glaser, R. (1991). The maturing of the relationship between the science of learning and cognition and educational practice. Learning and Instruction, 1, 129–144.

Haertel, E. H. (1990). Performance tests, simulations and other methods. In J. Millman & L. Darling–Hammond (Eds.), The new handbook of teacher evaluation. London: Sage.

Haertel, E. H. (1991). New forms of teacher assessment. In G. Grant (Ed.), Review of Research in Education, 17. Washington: AERA.

Mehrens, W. A. (1992). Using performance assessment for accountability purposes. Educational Measurement: Issues and Practice, 11, 1, 3–9.

Needels, M. C. (1991). Comparison of student, first–year, and experienced teachers' interpretations of a first–grade lesson. Teaching and teacher education, 7, 269–278.

Peterson, P. L., & Comeaux, M. A. (1989). Evaluating the systems: Teachers' perspectives on teacher evaluation. Paper presented at the conference of the American Educational Research association. San Francisco.

Powney, J. (1991). Teacher appraisal: The case for a developmental approach. Educational Research, 33, 83–92.

Sabers, D. S., Cushing, K. S., & Berliner, D. C. (1991). Differences among teachers in a task characterized by simultaneity, multidimensionality and immediacy. American Educational Research Journal, 28, 63–88.

Wilson, J. D. (1989). Introduction: Four issues in teacher assessment. In: J.D. Wilson, G. Thomson, R. Millward & T. Keenan (Eds.), Assessment for teacher development. London: The Falmer Press.

AN OVERVIEW OF TEN YEARS OF RESEARCH IN THREE DUTCH RESEARCH PROGRAMMES ON TEACHING AND TEACHER EDUCATION

F. Korthagen, & Th. Wubbels, IVLOS Institute of Education, University of Utrecht, Utrecht; B. van Hout-Wolters, Graduate School of Teaching and Learning, University of Amsterdam, Amsterdam, The Netherlands

Abstract

This symposium brings together the results of ten years of research in three Dutch research programmes on teaching and teacher education. The Nijmegen programme focuses on problem situations of teachers, their personality characteristics and their functioning in their professional practice, the Leiden programme on the content and structure of declarative knowledge of teachers and the Utrecht programme on educational arrangements in teacher education, with an emphasis on the promotion of student teachers' reflection. In each programme the development of adequate research techniques in the field of teaching and teacher education is a key issue.

Introduction

The Dutch system for the funding of university research changed ten years ago. A system was introduced in which research programmes are evaluated externally every five years. Funding is now dependent on these evaluations. Research programmes are required to consist of at least the equivalent of five full time researchers. They should have a coherent theoretical framework, an integrated plan of research activities and a coordinated research strategy.

It seems that this development fostered coherence in the Dutch research on teaching and teacher education. It also strengthened the quality of the research projects. Whereas before the eighties international publications of Dutch researchers in the field of teaching and teacher education were rare, reports of Dutch research projects now appear in international journals frequently.

The new system of requirements for university research has functioned during one decade. It is interesting to look back on the results of these ten years of research on teaching and teacher education within this new setting.

The symposium brings together researchers form three Dutch research programmes in this field. It gives an overview of the research projects in these programmes, their results and their dilemmas. Each presenter will also give an

outline of the key problems in the field and discuss the research strategies that were used as well as promising directions for future research.

THE INFLUENCE OF PERSONALITY CHARACTERISTICS ON TEACHERS' FUNCTIONING IN PROBLEMATIC SITUATIONS

Th. Bergen, & J. Giesbers, University of Nijmegen, Nijmegen, The Netherlands

The way in which teachers function during their professional practice is determined and influenced by the interaction of situational characteristics and personality characteristics. If this interactionistic starting point is seriously taken into account, both components of the interaction should be investigated.
This contribution reports about a research programme on problem situations of teachers, their personality characteristics and their functioning during their professional practice.
The results indicate that teachers experience professional situations as problematic when the problem is related to the core task of their profession viz. teaching and guiding pupils, when the core task is caused by other persons and when teachers cannot solve these problems easily. Problematic pupil-teacher interactions have these characteristics and consequently are important for teachers' functioning.
The way in which teachers perceive their professional problems is influenced by their beliefs and expectations concerning their influence on these problems. Especially personality characteristics such as locus, efficacy and attribution styles influence complaints with respect to helplessness, stress, work satisfaction and reaction styles.
The conclusion of this research is that psychological variables strongly influence teachers' professional practice.

DECLARATIVE KNOWLEDGE OF PROSPECTIVE PRIMARY SCHOOL TEACHERS

F. Boei, A.H. Corporaal, & F.K. Kieviet, University of Leiden, Leiden, The Netherlands

In this contribution the results are discussed of the research project "Cognitions of student teachers". This project can be divided into three phases. Central theme in all phases was the development in the declarative knowledge of student teachers concerning 'good education'.
In the first phase a large scale study was conducted with the aid of Kelly's repertory grid technique. 57 First-year student teachers and 60 third-year student teachers were interviewed with this technique. The results indicated that cognitions of student teachers can be described with the help of four central dimensions. Differences between first- and third-year students occurred only with respect to the dimension 'individual versus group'.

The second phase involved a more methodological concern. It appeared that data gathered with the repertory grid were of an inconsistent nature. Moreover, insight was only gained into the content of the declarative knowledge and not into the structure of this knowledge. The result of the second phase was a new instrument in which the respondents had to formulate the relations that, to their opinion, existed between the concepts they formulated and some given stimuli.
The third phase of the project concerned a small scale study into the substantial as well as structural development of declarative knowledge of first-year student teachers. The results indicated a move towards the concept of teaching, whereas a structural move was detected towards strategic knowledge (the declarative counterpart of procedural knowledge).

EDUCATIONAL ARRANGEMENTS IN PRESERVICE TEACHER EDUCATION

F.A.J. Korthagen, & J.S. ten Brinke, University of Utrecht, Utrecht, The Netherlands

A metastudy was conducted into five dissertation projects in the research programme on teacher education in the University of Utrecht. The projects focussed on educational arrangements in teacher education programmes. The combined outcomes of these projects are discussed. The projects are:

- The promotion of reflection in teacher education;
- "1:1" = student teaching in a setting with 1 student teacher and 1 pupil;
- An arrangement aiming at enhancing explorative competencies;
- A 4-6 months university teacher education programme aiming at the integration of theory and practice;
- The independent (Final) School-Practice Period in a 1-year postgraduate programme.

An important outcome is the fact that a common orientation underlying the arrangements could be identified and described, with the following characteristics:

- central aims for teacher education programmes are starting competence and competence for professional growth;
- learning to become a teacher is conceptualized as experiental learning;
- subjective theories of student teachers are the starting point for developmental processes;
- emphasis on self-evaluation by student teachers;
- gradual increase of complexity and responsibility in the teacher education programme;
- the contexts for student teaching are considered to be powerful factors;
- specific competencies are required from teacher educators.

A common characteristic of all arrangements is a strong theory-practice relationship. Project 4 shows that this relationship produced lasting effects in the teacher's starting competence, but not in the field of maintaining discipline. Project 5 indicates that the Independent Final School-Practice Period contributes to the competency of taking responsibility for teaching.

With regard to the promotion of reflection in teacher education, the studies show effects on various measures. For example, in a quasi-experimental study significant effects were found on interpersonal behaviour in the classroom and work satisfaction. It was also shown that learning orientations of student teachers appear to interact with the development of reflection.

963

INTERPERSONAL TEACHER BEHAVIOUR

N. Verloop, Teacher Education Department, University of Leiden, Leiden, The Netherlands

Introduction

In this symposium three research studies will be presented about teacher communication styles in relation to other aspects of the learning environment: teacher nonverbal behaviour, teacher instruction and school level environment characteristics. The studies fit in a larger research project on interpersonal teacher behaviour. In this project interpersonal teacher behaviour is analyzed from a systems perspective (Watzlawick, 1967) that has been adapted for use in classroom communication research by Wubbels, Créton and Holvast (1988). In this perspective it is assumed, that behaviours of the participants in communication influence each other mutually. Circular communication processes develop that not only consist of behaviour, but determine behaviour as well. Créton and Wubbels (1984) developed on the basis of work by Leary (1957) a model that maps interpersonal teacher behaviour on a proximity dimension (cooperation versus opposition) and an influence dimension (dominance versus submission). These two dimensions have been shown to be universal dimensions apt to describe interpersonal behaviour (Lonner, 1980). Different combinations of the two dimensions give eight different categories of teacher behaviours. Data about students' (and teachers') perceptions of teacher behaviour can be gathered with help of the Questionnaire on Teacher Interaction (QTI). This instrument is based on the model of interpersonal teacher behaviour and consists of eight scales, measuring the eight different aspects of teacher behaviour. Both the Dutch and the American version of this questionnaire have proven to be valid and reliable instruments to measure students' perceptions of relatively stable patterns in interpersonal teacher behaviour (Wubbels & Levy, 1991). With data obtained with this questionnaire, Brekelmans (1989) developed a typology of relatively stable patterns in interpersonal teacher behaviour. These patterns can be characterised as communication styles (e.g. a directive, an authoritative, an uncertain-tolerant, an uncertain-aggressive style). All the data about teachers gathered until now with the QTI in The Netherlands and in the US can be classified in the categories of the typology (Wubbels & Levy, 1991).

The papers in the symposium will extend this line of research by investigating if there is a relation between teacher communication styles and nonverbal behaviours of teachers (first paper), between teacher communication styles and

characteristics of the teacher's instruction (second paper) and between teacher communication styles and school environment (third paper).

TEACHER NONVERBAL BEHAVIOUR AND TEACHER COMMUNICATION STYLES

J. Van Tartwijk, M. Brekelmans, & H. Créton, University of Utrecht, Utrecht, The Netherlands

In this paper we will present results of a study on the relations between teacher nonverbal behaviour and students' perceptions of interpersonal teacher behaviour in the classroom. The aim of this study is to contribute to the knowledge base for the development of teacher education programmes to help teachers build productive communication styles.

The study focused on nonverbal behaviour because nonverbal behaviour seems the most important aspect of interpersonal behaviour (e.g. Mehrabian, 1981). To obtain data about the nonverbal teacher behaviour an observation instrument was developed, consisting of 21 observation categories together describing use of the available space in the classroom, distance from students, body orientation, body posture and body movement, facial expression, visual behaviour and some speech characteristics. According to the inter-observer agreement (index kappa) the instrument is reliable.

With the observation instrument we coded videotaped nonverbal teacher behaviour in time sampled fragments of one lesson of 27 student-teachers and of 8 lessons in the same class of 5 student teachers during a three months period (the first six lessons in that class and two lessons after some period of time). The student teachers were independent teachers for three months in secondary schools as a part of the Teacher Training program of the University of Utrecht (Koetsier, 1991).

With the QTI we gathered data on the students' perception of the teacher communication styles at the end of the three months period.

The first analyses show significant relations between the teacher nonverbal behaviour and students' perception of teacher communication styles on both the proximity and the influence dimension. Further analyses will focus on the longitudinal data on nonverbal behaviour. It will be investigated if characteristics of nonverbal behaviour during the first lessons in a new class can account for specific styles in the communication of student teachers after three months.

INSTRUCTIONAL EFFECTIVENESS AND INTERPERSONAL TEACHER BEHAVIOUR

J. Levy, & R. Rodriguez, George Mason University, Fairfax, United States of America; Th. Wubbels, University of Utrecht, Utrecht, The Netherlands

In previous research relations between the teacher communication style and student outcomes have been investigated (B.ekelmans, 1989). Strong relations were found between the dominance dimension and cognitive student outcomes. In research on teacher effectiveness characteristics of effective instruction have been identified (e.g. Brophy & Good, 1986). The results of these two lines of results raise the question if teacher instruction and teacher communication styles are related to each other. Are the results from these two lines of research simply equivalent? Do teachers who realise effective instruction at the same time realise effective communication styles? Or are the characteristics of the teacher's instruction and communication (partly) independent?
In this study therefore data were gathered on the teacher instruction with an observation instrument and about the teacher communication style with the QTI. The observation instrument, developed on the basis of a review of the literature on teacher effectiveness, consists of 23 items to be rated by an observer on a five point scale. The items can be grouped in six scales, measuring several aspects of teacher instructional effectiveness. The internal consistency of the scales was higher than 0.75. Inter-rater reliability was established on the basis of 14 lessons observed by two raters. The results showed acceptable reliability for all six scales. For the main part of the study data were gathered in 28 classes of teachers in secondary education in the US. Students answered the QTI once and an observer scored the teacher's lesson on the six observation scales on two occasions. The mean of the two observations was used as the measure for the teacher instruction. The data were analyzed with correlational analyses. It appears that teacher instructional effectiveness is correlated to the influence dimension of the QTI, but not to the proximity dimension. The correlations, however are not quite strong.

TEACHER COMMUNICATION STYLE AND SCHOOL ENVIRONMENT

D. Fisher, University of Tasmania, Launceston, Australia; B. Fraser, Curtin University, Perth, Australia; Th. Wubbels, University of Utrecht, Utrecht, The Netherlands

A useful distinction in learning environment research can be drawn between school-level and classroom-level environment. Whereas classroom climate involves relationships between teachers and their students or among students (Fraser, 1991), school climate might involve a teacher's relationships with other teachers, senior staff and the school principal. Over the last decade or so, the concept of school environment has appeared in the educational literature with

increasing frequency. In fact some studies in the literature suggest that the school environment makes a major contribution to the effectiveness of a school. The empirical evidence for this claim is not yet convincing (Scheerens, 1992). In any case it is likely that the study of the school environment contributes to understanding and improvement of the school's functioning and to satisfaction and productivity within schools. Learning environment researchers have however mostly focused in their studies on the classroom-level environment (Fraser, 1991). The present study brings this tradition together with the study of school-level environments.

In the framework of the research on interpersonal teacher behaviour presented in this symposium the question arose if the teacher communication style in the classroom is in any way influenced by or influences the school-level environment. The study involved the use of the School-Level Environment Questionnaire (SLEQ, Fisher and Fraser, 1990) to assess teachers' perceptions of school climate (scales student support, affiliation, professional interest, staff freedom, participatory decision making, innovation and resource adequacy) and the use of the Questionnaire on Teacher Interaction (QTI) to assess students' perceptions of teacher communication styles. Data were gathered in Australian classrooms in secondary education from both students (n = 969) and teachers (n = 46). Preliminary results show that in Australian classrooms the same types of teacher communication styles are found as in the Netherlands and the US. The analyses about the relations between the two types of environment are still under way and will be discussed in the paper.

References

Brophy, J. E. & Good, T. L. (1986). Teacher behaviour and student achievement. In M. C. Wittrock (Ed.), Handbook of research on teaching, (3rd ed.), pp. 328-375. New York: Macmillan.

Brekelmans, M. (1989). Interpersonal teacher behaviour in the classroom. (In Dutch): Interpersoonlijk gedrag van docenten in de klas. Utrecht: WCC.

Créton, H. A., & Wubbels, T. (1984). Discipline problems of beginning teachers. (In Dutch: Ordeproblemen bij beginnende leraren.). Utrecht: WCC.

Fisher, D. L., & Fraser, B. J. (1990, April). Validity and use of the School-level Environment Questionnaire. Paper presented at the Annual Meeting of the American Educational Research Association, Boston.

Fraser, B. J. (1991). Two decades of Classroom Environment Research. In B.J . Fraser & H.J. Walberg (Eds.), Educational Environments, (pp. 3-28). London: Pergamon Press.

Koetsier, C. P. (1991). Een brug tussen opleiding en praktijk.: Een studie over de individuele eindstage als brug tussen lerarenopleiding en praktijk. (Bridging the gap between teacher education and teacher practice). Dissertatie. Utrecht: WCC.

Leary, T. (1957). An interpersonal diagnosis of personality. New York: Ronald Press.

Lonner, W. J. (1980). The search for psychological universals. In H. C. Triandis & W. W. Lambert (Eds.), Handbook of Cross Cultural Psychology, (Vol.I.), pp. 143-204. Boston: Allyn and Bacon.

Mehrabian, A. (1981). Silent Messages. Belmont: Wadsworth.

Scheerens, J. (1992). Effective Schooling. London: Cassell.
Watzlawick, P., Beavin, J. H., & Jackson, D. (1967). The pragmatics of human communication. New York: Norton.
Wubbels, T., Créton, H. A.,& Holvast, A. J. C. D. (1988). Undesirable classroom situations. Interchange, 19, 2, 25-40.
Wubbels, T. & Levy, J. (1991). A comparison of interpersonal behaviour of Dutch and American teachers. International Journal of Intercultural Relationships, 15, 1-18.

968

PERSPECTIVES ON KNOWLEDGE BASES FOR TEACHING

Neville Bennett, Clive Carré, & Elisabeth Dunne, School of Education, University of Exeter, United Kingdom

Recent research on teaching from a constructivist perspective has focussed on the nature and quality of classroom tasks, and on the manner of their representation, and on the diagnosis of childrens' conceptions. This approach has taken due account of the role of the learner in mediating and structuring knowledge, and of the teachers level, and use, of appropriate knowledge bases for teaching. Thus the emphasis in research has shifted away from what teachers do, to what knowledge teachers hold, and how they organise that knowledge for teaching.

Central to this shift has been Shulman's (1987) identification of areas of teacher knowledge each with hypothesized, but by no means clear, a priori, relationships to teacher action. Research undertaken from this perspective has emphasised the importance of subject matter knowledge, both substantive and syntactic, in enabling teachers' to pose questions, to explain concepts, to select tasks, to evaluate their pupils understandings, and to make curricular choices (e.g. McDiarmid, Ball and Anderson 1989; and Grossman, Wilson and Shulman, 1989). However this research has tended to be on small samples of American student teachers, and largely at the secondary level.

Of concern in relation to the improvement of teacher education programmes, is the increasing evidence that teacher candidates have only a limited understanding of their subject (even when they have majored in it) and that many of the teacher education programmes studied appear to have little effect on the conceptions students have of their subject on entry (Kennedy 1991). Similarly there is evidence that students' incoming beliefs and attitudes about teaching and educational issues are limited and also difficult to change.

These findings raise questions about the precise nature of student teachers' entering knowledge and dispositions, the extent to which they develop and change, and in their relationship to teaching competence. Providing answers to such questions constituted the aims of the studies to be reported.

The research programme, the first of its type to be undertaken in Britain, addressed those questions in the context of the postgraduate training of elementary school teachers specialising in four subject areas - Maths, Science, English and Music.

The research focus was on the following areas:

- The knowledge and beliefs that students enter teacher training with;
- The impact on those of the processes of training, including both course structures/content and teaching practices;
- The relationship of subject knowledge to teaching competences.

969

After an introduction by the chairperson which will outline the appropriate theoretical perspectives and research methodology there will be three substantive papers. Three, rather than four, papers have been chosen deliberately to allow maximum opportunity for discussion. These are:

- The development of subject matter knowledge for teaching: Clive Carré;
- The processes of learning to teach: the development of knowledge for teaching: Elisabeth Dunne;
- The relationship of student knowledge and teaching competence: Neville Bennett.

THE DEVELOPMENT OF SUBJECT MATTER KNOWLEDGE FOR TEACHING

Clive Carré, School of Education, University of Exeter, United Kingdom

There is evidence to suggest that teacher candidates on post graduate courses of teacher training have only limited understanding of their subject. In order to ascertain students understanding of the subject matter that they would have to teach, assessments were developed based on the first six levels of the British national curriculum i.e. the level to which an able 11 year old, or average 14 year old, should be capable of achieving, in the core areas of English, Mathematics, Science and, additionally, Music. These included assessments of both substantive and syntactic knowledge. Each instrument thus covered knowledge of facts, concepts, relationships between conceptual ideas and methods of inquiry in the discipline. Instruments were also developed to assess students' beliefs about the nature of these subjects.

The entire cohort of students (N = 60) entering four specialist primary training courses were assessed during the first week of the course, and again at the end of the course. Analyses will present entering patterns of subject knowledge in each subject area, and the changes which occurred over the period of the training course. Areas of common difficulty will be analysed, and changes in subject knowledge, and attitudes toward the subjects will be evaluated in relation to academic qualifications, course content and student evaluations.

THE PROCESSES OF LEARNING: THE DEVELOPMENT OF KNOWLEDGE FOR TEACHING

Elisabeth Dunne, School of Education, University of Exeter, United Kingdom

Considerations of the relationships between teaching processes and student outcomes is common in classroom research but is almost non-existent in studies of learning to teach. Yet such relationships are critical to an understanding of the development of knowledge for teaching. The specific questions addressed here include the extent to which, and how, knowledge-for-teaching develops; the role in

this development of course structures and content, and the role of teaching practices.
In order to answer such questions several data gathering formats were used. In relation to taught courses student teachers kept structured diaries to record course input and their reflections and evaluations of their value. These were contextualised by interviews with tutors to provide course aims, intentions and content.
In relation to teaching practices students agreed to keep a separate diary focussing particularly on the role of the class teacher and the university supervising tutor. Diaries were supplemented by interviews with the students, teachers and tutors to gain clear views of the different participants perspectives.
Data will be presented on the impact of input on student learning, particularly in the areas of pedagogic content knowledge and knowledge of learners and curriculum. Data will also be presented on the roles which successive teaching practices played in student development, and the roles of the classteacher and supervising tutor in that development.

THE RELATIONSHIP OF STUDENT KNOWLEDGE AND TEACHING COMPETENCE

Neville Bennett, School of Education, University of Exeter, United Kingdom

Studies of the relationship between student knowledge bases and teaching competence have led to the conclusion that teachers cannot teach what they do not know. However these studies have tended to be American in origin and based largely at secondary school level. Do these relationships necessarily hold in the training of more generalist elementary teachers? To what extent, for example, do mathematics graduates teach mathematics to higher levels of competence than non-mathematics graduates in the elementary classroom? Similarly, to what extent do mathematics graduates teach mathematics to higher levels of competence than they themselves teach other subjects?
Twelve students were sampled from the cohort to provide four students from each of three different, but overlapping primary training courses, one specialising in mathematics, one in science and one in music. Each was observing for four days during their teaching practices, ensuring a minimum of 10 lessons, some being their specialist subject and others not. For each lesson students completed a pre-lesson questionnaire setting out their aims, the tasks chosen, how these would be presented, how the classroom would be organised and how childrens learning would be assessed. The lesson itself was and also observed by a trained observer. The observer then carried out a post-lesson interview to acquire students evaluations and reflections.
The transcripts of the lessons, and the observations have been analysed using a category system developed around Shulman's knowledge bases, and competences derived from recent theoretical and empirical research on teaching and learning. Data will be presented in both quantitative and qualitative form to address the relationship between student knowledge bases and teaching competence.

AUTHORS INDEX

Name	Chapter	Page
Aalsvoort, G. M. van der	3	375
Aarnoutse, C. A. J.	3	295
Aarnoutse, C. A. J.	5	719
Achtenhagen, F.	7	945, 999, 1035, 1037
Ackerman, T. A.	5	705
Admiraal, W.	6	823
Aguilera, A.	6	908, 911
Aho, L.	2	172
Akker, J. van den	2	99, 153, 189
Akkök, F.	5	731
Allegrezza-Carvoyeur, L. S.	9	1251
Alvarez, P.	5	668
Amelsvoort, J. van	3	396
Anderson, C.	8	1130
Ansellem, M.	3	320
Antoli, B.	6	894
Antunes, H.	9	1243
Ardaç, D.	9	1268
Askar, P.	3	258
Askar, P.	5	731
Assink, E. M. H.	3	285
Assink, E. M. H.	4	644
Atkins, M. J.	4	531, 651
Babanin, L.	4	643
Barendse, P.	8	1199
Barker, P.	4	592
Barrett, E.	6	854
Barrows, H. S.	7	954
Barton, L.	6	854
Bassey, M.	2	213
Bates, I.	9	1241
Becher, T.	8	1063
Behrens, M.	7	1001
Beishuizen, J. J.	3	432, 436
Beishuizen, J. J.	4	633
Benins, A.	6	904
Bennett, N.	3	400
Bennett, N.	6	939, 941
Berberoglu, G.	5	708, 748
Berg, E. van den	2	146
Berg, G. van den	8	1158
Berg, J. van den	7	1020
Berg, R. van den	4	521
Bergen, J. van	5	786
Bergen, Th.	3	396, 396
Bergen, Th.	6	931
Berger, M. P. F.	5	692
Berger, N.	3	387
Bergh, H. van den	2	108
Bergmans, A.	3	295
Berings, I.	9	1286
Bestebreurtje, R.	4	584
Bijtel, J. A. L. van	6	796
Black-Branch, J. L.	9	1276
Blackman, S. J.	4	622
Blom, S.	1	79
Boei, F.	6	931
Boekaerts, M.	3	302, 360, 395
Boekenoogen, M. S.	7	994
Boekhorst, A. K.	4	619
Boekkooi-Timminga, E.	5	698
Bock, R. D.	5	662, 753
Bögels, S. M.	3	245
Bokhove, J.	5	783
Bollen, R.	1	71
Bolzoni, M.	4	569
Bonekamp, L. W. F.	4	538
Boon, J.	7	968
Börger, M. A.	6	924
Bosch, H. van den	8	1201
Bosch, L. van den	5	786
Bos, K. Tj.	1	63
Bos, K. Tj.	6	862
Bosker, R. J.	1	19, 86
Bosker, R. J.	9	1215
Bosman, Y.	7	1004
Bostock, S. J.	4	502
Bossink, C.	6	922
Bouhuijs, P. A. J.	8	1187, 1188
Braaksma, J.	4	577
Brandsma, H. P.	1	8, 69
Brandsma, H. P.	6	844
Brandsma, T. F.	7	1040, 1053
Bratko, D.	5	735
Brekelmans, M.	6	935
Breuer, K.	3	415
Brinke, D. ten	5	774
Brinke, J. S. ten	6	820, 932
Bristow, S.	3	405
Bristow, S.	9	1257
Brouwer, K.	7	964

Brown, A. 7 1001
Brown, A. 8 1168
Bruggencate, G. C. ten 8 1107
Bruijn, E. de 7 1015
Brummelhuis, A. ten 2 206
Bruinsma, C. 6 837
Büchel, F. P. 3 407
Büchel, P. 3 407, 409
Burton, R. M. 1 94
Busato, V. V. 3 435
Cachapuz, A. 6 816
Calderheads, J. C. 6 878
Cardelle-Elawar, M. 3 371
Cardelle-Elawar, M. 6 829
Carleer, G. J. 2 186
Carleer, G. J. 4 434, 631
Carré, C. 6 939, 940
Cate, Th. J. ten 6 796
Cate, Th. J. ten 8 1141
Cavendish, S. 9 1221
Cerveró, F. 6 894
Chartier, D. 3 329
Christiaans, H. H. C. M. 3 401
Claessens, M. P. J. G. 8 1196
Clarke, S. 7 1010
Clement, M. 6 885
Cohen, R. 4 587
Collis, B. A. 4 439, 534, 631
Collins, D. 4 502
Commins, B. 4 525
Confalonieri, E. 9 1245
Connelly, F. M. 2 213
Corporaal, A. H. 6 931
Corte, E. De 3 351, 393, 398, 417
Cox, M. 4 631
Creemers, B. P. M. 1 74, 82
Cremers, L. 4 545
Cremers, L. 6 862
Créton, H. 6 935
Cuijk, H. van 9 1286
Czerniawska, E. 3 345
Daal, M. M. 3 429
Daal, V. H. P. van 3 293
Daly, P. 1 60
Dam, G. T 9 1218
DeAyala, R. 5 686
Dekkers, H. 9 1226
Demirel, M. 3 342
Desforges, C. 3 405
Desforges, C. 9 1257
Diana, I. De 4 595, 598
Diana, I. De 5 728
Diggele, J. van 5 783
Dijk, H. van 3 263
Dijk, W. van 3 354
Dijkstra, S. 3 223, 414
Dirkzwager, A. 4 634
Dochy, F. J. R. C. 3 398, 417
Doets, C. 7 1043
Dolmans, D. 8 1133
Donald, J. G. 8 1145
Dong, Z-P. 4 557
D'Onofrio, A. 8 1172
Doornekamp, G. 4 545, 605
Doughty, P. L. 4 625
Draper, J. 6 813
Driel, C. van 6 820
Driel, J. van 8 1189
Dronkers, J. 9 1271
Dumitrescu, A. 2 131
Dumitrescu, A. 6 856
Dunne, E. 6 939, 940
Dunne, R. 6 871
Eash, M. J. 2 169
Eck, E. v. 9 1218
Eberg, J. W. 2 111
Eberg, J. W. 7 994
Edelenbos, P. 1 69
Eeden, P. van den 2 100
Eggen, T. 5 786
Ehrlich, D. 4 572
Eijkelhof, H. M. C. 3 256
Eldering, L. 9 1282, 1283
Elen, J. 8 1178
Elliott, R. G. 6 834
Elshout, J. J. 3 310, 435, 461
Elshout-Mohr, M. 3 434
Elst, G. van 7 977
Ely, D. P. 4 480
Engelen, R. 5 786
Engelhard, G. 5 741
Erder, A. M. 3 342

Ersoy, Y. 4 483
Ersoy, Y. 6 896
Etelpelto, A. 7 951
Evans, D. 1 55
Fátima Choráo C. Sanches, M. de 3 269
Feijen, C. J. 7 1029
Fergusson-Hessler, M. 3 232
Fernandez Benitez, V. 6 858
Feteris, A. 7 1017
Finegold, M. 4 457
Fischer, D. 6 936
Fitz-Gibbon, C. T. 1 16
Fleming, R. 6 904
Forcheri, P. 3 323
Fox, M. 8 1172
Fraser, H. 6 813
Fraser, B. 6 936
Free, E. L. 8 1114
Frey, K. 2 197
Fricke, R. 4 617
Frijns, P. 8 1193, 1199
Furlong, J. 6 854
Gaag, F. J. van der 4 613
Geban, O. 3 258
Geest, T. M. van der 2 187
Gelderblom, A. 7 980
Geus, W. de 3 298
Geus, W. de 5 744
Gielen, I. 3 351
Giesbers, J. 1 90
Giesbers, J. 6 931
Gijselaers, W. H. 8 1133, 1202
Gillijns, P. 5 786
Glas, K. 5 774
Glendinning, A. 9 1254
Glopper, K. de 3 298
Gmelich Meijling, M. R. 4 490, 613
Goldstein, H. 1 11
Goldstein, H. 5 751
Gómez-Crespo, M. A. 3 381
Gorsky, P. 4 457
Graaff, E. de 8 1187, 1188, 1193, 1208
Grace, G. 6 828
Gravemeijer, K. 2 103
Gray, J. 1 40
Groen-de Jong, C. 3 364
Groot, M. de 3 375
Groot, T. 4 619
Groppo, M. 9 1245
Gruber, H. 7 1038
Grunefeld, H. 8 1069
Guillén, J. 3 390
Gulmans, J. 4 521
Haan, D. M. de 1 75
Haapala. A. 2 172
Hacquebard, A. E. N. 2 165
Halász, G. 9 1275
Hambleton, R. K. 5 681
Hamers, J. 3 407
Hameyer, U. 2 99, 153
Hansen, K-H. 4 487
Harlen, W. 2 175
Harvard, G. 6 871
Heathcote, G. 6 793
Heitmann, G. 8 1126
Hell, J. G. van 4 461
Helmke, A. 3 398
Helsper, E. 4 640
Hendry, L. B. 9 1254
Henneken, O. W. H. 5 744
Hermans, H. 3 422
Herrlitz, W. 2 212
Heuvel-Panhuizen, M. van den 2 103
Heuvel-Panhuizen, M. van den 3 312
Hippe, Z. S. 4 465
Hoeben, W. Th. J. G. 2 210, 212
Hoeven-van Doornum, A. A. 9 1264
Hofman, R. H. 1 46
Holmberg, L. 5 717
Holsbrink-Engels, G. A. 3 243
Hoogstraten, J. 8 1077
Hopkins, D. 1 62
Hough. M. 6 846
Hout, J. F. M. J. van 8 1155
Houtveen, Th. 6 802
Hout-Wolters, B. H. A. M. van

	3	432
Hout-Wolters, B. H. A. M. van	6	791, 930
Hövels, B.	7	1020
Huijsman, Z. C.	8	1081, 1153
Hulst, A. van der	3	357
Hutchison, D.	5	756
Huteau, M.	3	329
Ibrahim, A. B.	1	16
Izard, J.	5	775
Jagt, P. van der	6	837
Jahn, H.	8	1123
Jakowicka, M.	3	348
James, C.	6	878
Jansen, E. P. W. A.	8	1073
Janssen, J.	5	783
Janssen, J.	9	1228
Janssen, P. J.	3	402
Janssen, T.	7	1041
Janssen Reinen, I.	2	202
Järvelä, S.	3	272
Jaspers, J. G. M.	4	560
Jeffcut, P.	8	1081
Jenkins, H.	1	55
Jochems, W. M. G.	8	1084, 1088, 1196, 206
Johnson, S.	1	13
Jong, F. P. C. M. de	3	396
Jong, J. de	5	721
Jong, J. A. de	7	971, 977
Jong, M.J. de	6	865
Jong, R. de	6	806
Jong, W. J. de	8	1088
Joosten, G.	9	1228
Jorge, M.	6	816
Jost, K. L.	4	625
Jung, M. H.	4	617
Jungbluth, P.	9	1264, 1284, 1311
Jurgens, H. J. H.	8	1155
Kaldeway, J.	3	337
Kamp, L. J. Th. van der	5	759
Kamp, M. van der	7	1041
Kanselaar, G.	2	218
Kanselaar, G.	4	560
Karstanje, P. N.	1	42
Karstanje, P. N.	7	1004
Kattenberg, G.	3	285
Kauppi, A.	6	905
Kessels, J.	7	1025
Kesteren, B. J. van	4	534
Keursten, P.	2	143
Kieviet, F.	6	919, 931
Kirschner, P. A.	3	417, 422
Kleintjes, F. G. M.	1	25
Kley, P. van der	9	1286
Kluvers, C.	3	396
Knuver, A. W. M.	1	8
Koetsier, C. P.	6	820
Kohlemainen, P.	8	1111
Kohn, D. I.	3	384
Kok, W. A. M.	3	306
Kok, W. A. M.	4	560
Kommers, P.	4	580
Koning, J. de	5	739
Koning, J. de	7	980
Kontogiannopoulou-Polydorides, G.	2	198
Koolstra, C.M.	5	759
Koopman, P.	8	1100
Kop, J. L.	9	1251
Koper, E. J. R.	4	634
Korthagen, F.	3	337
Korthagen, F.	6	823, 930, 932
Korthals, M.	8	1182
Kovacevic, P.	5	735
Kraemer, J. M.	5	783
Krause, B.	3	267
Kremers, E. J. J.	1	25
Krogt, F. J. van der	7	973
Krykorková, H.	3	249
Kuiper, W. A. J. M.	2	155
Kusá, D.	6	916
Ladhani, A-N.	4	595
Landucci, M.	2	117
Lang, M.	2	200
Langereis, M.	8	1077
Lautrey, J.	3	329
Leal Filho, W. D. S.	2	121
Lebedev, A. N.	5	734
Leemkuil, H. H.	3	325
Leenders, Y.	6	900
Leeuwe, J. F. J. van	5	719
Leiblum, M. D.	4	549
Lera, M.	6	908

Letiche, H. K.	8	1081, 1153
Levy, F.	3	320
Levy, J.	6	936
Lewis, B.	4	468
Lijnse, P. L.	3	256
Lima, A. de	6	880
Linden, E. van der	4	567
Linden, J. van der	4	644
Linden, W. van de	5	657
Lindström, B.	4	632
Lingefjärd, T.	4	647
Liverta Sempio, O.	9	1245
Loarer, E.	3	329
Lockwood, F.	5	660
Lodder, B. J. H.	8	1057
Lodder, B. J. H.	9	1235
Logar-Djuric, S.	5	735
Lombardini, H. E.	2	117
Loo, P. J. E. van de	8	1057
Lopez, A. M.	5	762
Love, J. G.	9	1254
Lowyck, J.	3	415
Lowyck, J.	6	840
Luyten, H.	1	74
Macready, G.B.	5	686
Main, R. G.	3	368
Maltseva, I.V.	5	734
Mandl, H.	7	1038
Manfred, E.	5	685
Mantovani, G.	4	569
Mª Poulussen van Balen, J.	6	858
Marinov, S.	5	711
Markinev, A. V.	5	734
Martens, R. L.	3	429
Martinez, R.	9	1293
Martinkova, V.	2	128
Martiny, U.	8	1060
Mason, G.	6	936
Mason, L.	3	333
Masters, G. N.	5	664
Matthijsen, C.	6	844
McGuigan, L.	2	209
McGuigan, L.	3	261
Meer, L. v. d.	8	1182
Meer, S. van der	8	1133
Meester, M. A. M.	3	422
Meesterberends-Harms, G. J.	7	958
Meijer, P. J.	3	281
Meijer, P. J.	4	563
Meijer, R. R	5	665
Mellink, E. C.	7	1004
Merriënboer, J. van	3	411
Menis, J.	2	213
Meulemann, H.	9	1273
Meyer, J. R.	6	849
Meyer, P.	6	900
Middelbeek, E.	3	422
Miles, S.	6	854
Milikowski, M.	3	310, 435
Miller, L.	4	609
Millett, A.	2	208
Min, F. B. M.	4	490, 613
Minnaert, A.	3	402
Mirande, M. J. A.	4	549
Moelands, F.	5	786
Moerkamp, T.	7	1044
Molenaar, I. W.	5	673
Molfino, M. T.	3	323
Mooij, T.	2	100
Moonen, J.	4	439, 633
Mora, J. G.	6	911
Mora, J. G.	9	1293
Morais, A. M.	9	1243
Morales, M.	5	762, 764
Moreno	5	764
Moskovkina, A. G.	9	1262
Mourik, T. G. C. van	3	245
Mulder, F.	2	165
Mulder, J. W.	8	1117
Mulder, L.	9	1312
Mulder, M.	7	945, 1021
Munn, P.	8	1093
Muraki, E.	5	662
Naylor, A.	6	914
Neber, H.	7	1039
Neut, I. v. d.	4	446
Neuwahl, N. M. E.	3	364
Nicaise, I.	9	1231
Niederer, R.	2	197
Niet, G.	4	613
Nieuwenhuis, A. F. M.	7	991, 1015
Nieuwenhuizen, C	3	302
Nieveen, N.	2	143
Nijhof, W. J.	7	945, 1033

Nguyen-Xuan, A. 3 320
Obel, B. 1 94
O'Brien, J. P. 6 914
O'Connor, J. E. 4 454
Oestermeier, U. 3 236
Ojen, Q. H. J. M. van 1 42
Ojen, Q. H. J. M. van 6 851
Ok, A. 6 809
Olsen, J. 2 184
Olsen, J. 4 609
Onna, B. van 7 960, 1029
Onstenk, J. 7 1046
Oort, F. J. 5 708, 748
Oosterhuis-Geers, J. 8 1065
Otten, R. 3 395
Oquist, O. 2 141
Oud, J. H. L. 5 719
Pan, H. 1 11
Pan, H. 5 751
Pavia, F. 6 894
Paykoç, F. 6 809
Peisa, S. 6 905
Pelgrims, G. 3 407, 410
Pelgrum, W. J. 1 75
Pelgrum, W. J. 2 195, 202
Perez-Eccheverria, M. P. 3 378
Petrikás, A. 1 31
Pieters, J. M. 6 791
Pieters, M. 2 115
Pijnenborgh, F. 3 250
Pilot, A. 4 577
Plomp, Tj. 2 193, 195, 202, 213
Plomp, Tj. 4 534
Pluijm, J. van der 9 1259
Podolskij, A. E. 8 1181
Poelmans, P. 3 429
Pool, I. 7 977
Poorthuis, G. M.T. 3 306
Popham, W. J. 5 713
Poschner, M. T. 8 1165
Pouw, C. L. M. 8 1114, 1182
Powney, J. 2 213
Powney, J. 8 1161
Pozo, J. I. 3 378, 381
Prucha, J. 2 125
Puchová, I. 3 276
Qualter, A. 2 209
Raban, B. 2 208
Raffe, D. 7 985
Ramaekers, G. W. M. 8 1057
Rasbash, J. 1 11
Rasbash, J. 5 751
Reezigt, G. J. 1 34
Reezigt, G. J. 9 1314
Reimerink, H. B. 4 613
Reitsma, P. 3 289, 298
Reitsma, P. 5 744
Reitsma, P. 6 874
Reisby, K. 2 135
Renkl, A. 7 1038
Renyier, M. 2 117
Resing, W. C. M. 3 387
Reynolds, D. 1 62
Reynolds, L. 4 572
Rezendes, G. 5 685
Riet, K. van 't 4 493
Rijns, Y. 8 1088
Riley, K. A. 1 58
Ris, Th. P. J. 4 615
Robijns, M. C. 9 1279
Roebuck, M. 4 529
Roeleveld, J. 9 1279
Rodriguez, D. 3 390
Rodriguez, R. 6 936
Roes, M. G. 2 148
Rojero, F. F. 2 117
Roode, F. de 5 774
Ropo, E. 2 133
Ropo, E. 6 826
Ruijs, G. P. C. M. 7 1050
Russell, T. 2 206, 209
Russell, T. 3 261
Sagdullaev, A. A. 9 1262
Sainsbury, M. 5 779
Saldaña, D. 3 390
Salonen, U. 2 172
Samarapungavan, A. 3 435
Sanders, M. 6 900
Sanz, A. 3 378
Saroyan, A. 8 1145
Schadé, E. 8 1141
Schagen, I. 5 756, 776
Schaick, P. van 4 513

Schaick Zillesen, P. G. van 4 490, 493, 613
Scheerens, J. 1 3, 80, 85, 86
Schmidt, H. 8 1133
Schmidt-Hackenberg, B. 7 989
Schmitt, M. 3 285
Schoenmaker, J. 4 468
Schott, F. 3 416
Schott, F. 8 1181
Schouten, A. 4 637
Schranhorst, U. 3 407, 411
Schreurs, M. L. J. J. 7 947
Schuurs, U. 2 108
Schuurs, U. 5 721
Sciarone, A. G. 3 281
Sciarone, A. G. 4 563
Scaratti, G. 9 1245
Seegers, G. 3 302, 360
Segers, S.R. 5 772
Sijtsma, K. 5 665
Sijtstra, J. 5 784
Simons, P. R. J. 3 395, 396, 411, 417
Simons, R-J. 3 223
Siregar, N. 6 867
Sheridan, B. E. 5 671
Shucksmith, J. 9 1254
Slettenhaar, H. K. 8 1117
Sliepen, S. E. 6 874
Sligte, H. W. 4 509
Smeets, D. 2 165
Smid, H. J. 8 1088
Smith, K. 5 772
Snippe, J. 2 151
Sola, I. 3 390
Spencer, E. P. 1 5
Spoon, P. 4 446
Squarcy, K. 5 735
Staa, A. van 7 968
Staessens, K. 6 891
Stam, H. van der 8 1120
Stanchev, I. 4 439
Stanley, S. J. 2 138
Steiner, G. 3 239
Stenström, M. L. 9 1248
Stoel, W. G. R 6 844
Stokking, K. M. 1 49
Stolk, H. 3 425
Stoutjesdijk, E. 3 436
Streumer, J. N. 7 945, 1009, 1017
Strum, J. 2 212
Struyf, E. 3 351
Sultana, R. G. 9 1238
Sumner, R. 8 1102
Taconis, R. 3 232
Tartwijk, J. van 6 935
Taylor, W. 6 813
Terekhina, A. 4 641
Tergan, S-O. 3 236
Terlouw, C. 8 1114, 1175
Terwel, J. 2 99, 100
Thompson, J. 4 525
Tiemessen, W. J. M. 8 1117
Tillema, H. H. 6 919
Timmer, E. 8 1091
Timmer, J. 4 544
Tobi, H. 5 695
Toluk, Z. 6 896
Tomic, W. 2 168
Trumper, R. 2 225
Tuijnman, A. 2 206
Tymms, P. B. 1 37
Ulbricht, K. R. G. 5 767
Underwood, G. 9 1221
Underwood, J. 9 1221
Valcke, M. M. A. 3 429
Valk, A. E. van der 3 256
Vandenberghe, R. 6 882, 885, 888, 891
Vedder, P. 9 1282, 1283
Veen, A. M. 9 1279
Veen, C. H. J. van 8 1081, 1153
Veen, J. T. van de 8 1107
Veen, R. van de 8 1091
Veen, W. van 4 446
Veenman, M. V. J. 3 435, 461
Veenman, S. 6 900
Veenstra-Strijland, T. 2 165
Vegt, A. L. van der 6 802
Velden, L. v. d. 9 1318
Velden, R. van de 9 1235
Velden, R. K. W. van der 8 1057
Vendrig, A. S. A. 7 971

Verhagen, P. W.	4	518, 584
Verloop, N.	6	926, 934
Vermeulen, C.	1	19
Vermeulen, C.	6	862
Vermeulen, M.	1	42
Vermeulen, M.	8	1096
Vermunt, J. D. H. M.	7	983
Vermunt, J. D. H. M.	8	1079
Verschaffel, E.	3	351
Versloot, A. M.	7	971
Verwey, A.	8	1088
Vila, L.	9	1293
Vinjé, M. P.	5	785
Vinke, D.	8	1084
Visscher, A. J.	1	63
Visser, H. J.	1	42
Vlas, H.	4	545
Voeten, M. J. M.	3	295
Voeten, M. J. M.	9	1264
Vogelzang, F.	4	446
Volman, M.	9	1218
Volpe, R.	6	904
Volman, C.	5	728
Voncken, E.	7	1004
Voogt, J. M.	2	184, 188
Vooijs, M. W.	5	759
Voronel, Y.	9	1289
Vos, H.	3	228
Vos, H. J.	5	725
Vos, M.	1	28
Vouten, M. J. M.	5	719
Vries, S. de	4	580
Vrugt, A.	8	1077
Vu, N. V.	7	954
Walker, D. F.	2	163
Walraven, M.	3	289
Walsh, A.	2	208
Wang, J. W.	4	541
Warmerdam, J.	7	973
Watson, D. M.	4	441
Weerden, J. van	5	783
Weide, M. G.	1	34
Weide, M. G.	9	1315
Werf, G. van der	9	1296
Wesselingh, A.	9	1215, 1286
Westbury, I.	2	182
Weston, P.	2	179
Wetterling, J.	4	553
Wieringen, F. van	1	88
Wiggers, J. A.	7	991
Wierstra, R. F. A.	2	159
Wijngaarden, K. van	8	1138
Wilbrink, B.	5	701
Wilbrink, B.	8	1149
Wignall, R.	1	53, 67
Willems, J.	8	1138
Wise, S. L.	5	689
White, J.	6	800
Whitty, G.	6	854
Witziers, B.	1	21
Woodhouse, G.	1	11
Woodhouse, G.	5	751
Woord, J. van der	8	1208
Woude, J. van der	4	506
Wubbels, Th.	6	820, 823, 930, 936
Yang, M.	1	11
Yang, M.	5	751
Yildiz, R.	4	651
Yontar, A.	3	253
Yrjönsuuri, R.	3	317
Zarevski, P.	5	735
Zeeuw, G. de	4	509
Zhang, J-P.	4	557, 602
Zhu, Z.	4	598
Zimowski, M. F.	5	753
Zinn, B. R.	3	384
Zinn, B. R.	9	1289
Zolingen, S. J. van	7	960
Zuidema, J.	4	640
Zuilichem, J. van	7	983
Zysno, P. V.	5	680
Zwart, W-J.	4	498
Zwietering, M. H.	4	493

22 - 25
JUNE
1992

BOOK OF SUMMARIES

EUROPEAN

UNIVERSITY OF TWENTE

CONFERENCE

DEPARTMENT OF EDUCATION

ON

EDUCATIONAL

ENSCHEDE - THE NETHERLANDS

RESEARCH

981

BEST COPY AVAILABLE

EDITORS: TJ. PLOMP, J.M. PIETERS, A. FETER

European Conference on Educational Research

BOOK OF SUMMARIES

Volume 3

University of Twente
The Netherlands
June 22 - 25, 1992

982

COLOFON

Editors:
Tjeerd Plomp
Jules Pieters
Andries Feteris

Editorial assistance:
Harmen Abma
Jeroen Breman
Conny de Koning
Olivia Kramers
Renate Schraa

Cover:
Hanna Snijder

Print:
Duoprint

University of Twente
Department of Education
P.O. Box 217
7500 AE Enschede

ISBN 90-365-0536-4

Copyright © 1992 by Department of Education, University of Twente

983

CONTENTS

Chapter 7
VOCATIONAL AND PROFESSIONAL EDUCATION AND TRAINING 943

Papers
Introduction to theme 7: Vocational and professional education and training, *W.J. Nijhof, F. Achtenhagen, M. Mulder, & J.N. Streumer* 945
Career expectations of students of the health sciences, *M.L.J.J. Schreurs* 947
Work experience and the development of expertise, *A. Etelpelto* 951
Feasibility, reliability, and validity of using the standardized/simulated person technique in performance assessment of competence in the professions, *N.V. Vu, & H.S. Barrows* 954
Modular instruction, student investment and school output, *G.J. Meesterberends-Harms* 958
Key qualifications and intermediate vocational education (MBO), *S.J. van Zolingen, & B. van Onna* 960
Work simulations in Dutch adult vocational education centres, *K. Brouwer* 964
Effects of adult education: overview and perspectives of research in the Netherlands, *J. Boon, & A.van Staa* 968
Structured on-the-job training in the Netherlands, *J.A. d. Jong, A.S.A. Vendrig, & A.M. Versloot* 971
Qualification-policy in professional organisations: an empirical exploration of some theoretical propositions about learning - and qualification - strategies in different types of organisations, *F.J. v.d. Krogt, & J. Warmerdam* 973
Stacking and tinning: how steelworkers learn their jobs, *J.A. d. Jong, I. Pool, & G. Van Elst* 977
Company training: volume, underinvestment, and return, *J. d. Koning, & A. Gelderblom* 980
Learning and training styles in corporate training, *J. van Zuilichem, & J.D.H.M. Vermunt* 983
The 'new flexibility' in vocational education, *D. Raffe* 985
New methodologies for in-company vocational training and education: Empirical cases studies, *B. Schmidt-Hackenberg* 989
Apprenticeship models in the metal industry: figures and vision, *A.F.M. Nieuwenhuis, & J.A. Wiggers* 991
Evaluation of apprentice-system courses in information science, *M.S. Boekenoogen, & J.W. Eberg* 994
How should research on vocational and professional education react to the challenges in life and worksite, *F. Achtenhagen* 999
Labour market and opportunities of becoming skilled in England and the Federal Republic of Germany, *M. Behrens, & A. Brown* 1001
Assessment of the effectiveness of careers counsellors in secondary education, *E. Voncken, E.C. Mellink, Y. Bosma, & P.N. Karstanje* 1004

Symposia

Responsiveness and organisational effectiveness in vocational education and training, *J.N. Streumer* 1009

Issues in developing curriculum programs for corporate training and education, *M. Mulder* 1021

Knowledge utilisation within vocational education, *W. Nijhof* 1033

Education and training for lower educated adults, *J. Brandsma* 1040

Chapter 8

HIGHER EDUCATION 1055

Papers

Sex differences in the labour market position of young graduates, *B.J.H. Lodder, P.J.E. v.d. Loo, G.W.M. Ramaekers, & R.K.W. v.d. Velden* 1057

Paths of learning and professional integration women and their colleagues in architecture, *U. Martiny* 1060

Disciplinary cultures, *T. Becher* 1063

Procedure to promote effective and efficient study skills (PROPES), *J. Oosterhuis-Geers* 1065

Why some girls do drop out of technical studies and others don't, *H. Grunefeld* 1069

Educational programming related to study progress, *E.P.W.A. Jansen* 1073

The relation of perceived self-efficacy and goals to academic performance, *A. Vrugt, M. Langereis, & J. Hoogstraten* 1077

Constructive learning in higher education, *J.D.H.M. Vermunt* 1079

Escaping from mass produced higher education, *H.J. v. Veen, H.K. Letiche, Z.C. Huijsman, & P. Jeffcutt* 1081

Internalization of education: A foreign language as the medium of instruction, *D. Vinke, & W. Jochems* 1084

Academic success of foreign students in engineering education, *W. Jochems, W.J. d. Jong, Y. Rijns, H.J. Smid, & A. Verweij* 1088

Older students in higher education: An investigation into their wants and attitudes, *E. Timmer, & R. v.d.Veen* 1091

Adult access to higher education in Scotland: The Scottish wider access programme, *P. Munn* 1093

Regional differences in aggregate enrolment levels in higher education, *M. Vermeulen* 1096

Who stirs up the inflation of graduation in higher education?, *P. Koopman* 1100

What students expect to do after 16: The school's role, *R. Sumner* 1102

CAL on digital design: Student controlled practising & testing, *G.C. t. Bruggencate, & J.T. v.d. Veen* 1107

Computer anxiety at the beginning of the studies, *P. Kohlemainen* 1111

The transfer effect of teaching systematic problem solving techniques in computer assisted learning, *E.L. Free, C.L.M. Pouw, & C. Terlouw* 1114

Problem solving and sequencing for problem solving in mechanisms in higher education, *J.W. Mulder, W.J.M. Tiemessen, & H.K. Slettenhaar* 1117

Learning aid for problem solving in physics, *H. v.d. Stam* 1120

The implementation of the 'Fachhochschulen' in the new federal states of Germany, *H. Jahn* 1123
Expanding higher education structures: The type of 'Fachhochschulen' in the European context, *G. Heitmann* 1126
University students' perceptions of their experience of tutorial groups, *C. Anderson* 1130
Students' learning activities beyond teachers' objectives, *D. Dolmans, W. Gijselaers, &H. Schmidt* 1133
Directing learning activities by means of tasks in problem based learning, *K. v. Wijngaarden, & J. Willems* 1138
Preparing students for clinical problem solving in the pre-clinical curriculum, *Th.J. t. Cate, & E. Schadé* 1141
Evaluating teaching performance in Canadian universities, *A. Saroyan, & J.G. Donald* 1145
The first year examination as negotiation: An application of Coleman's (1990) social system theory to law education data, *B. Wilbrink* 1149
Selection of students during first year of higher education, *Z.C. Huijsman, C.H.J. v. Veen, & H.K. Letiche* 1153
Development and testing of a causal model for assessing problems of Dutch research assistants (PhD. students), *J.F.M.J. v. Hout, & H.J.H. Jurgens* 1155
A usable approach for evaluation in higher education, *G.v.d. Berg* 1158
Views on 'peer review' and quality assurance, *J. Powney* 1161
The educational research at the national university of Mexico, *M.T. Poschner* 1166

Posters
Financing of student status: Anglo-German comparisons, *A. Brown* 1169
The place of planning in leadership education, *A. D'Onofrio, & M. Fox* 1172

Symposia
The FUNDES-case: A model for the development of instructional design, *C. Terlouw* 1175
The development and implementation of a problem-based curriculum, *E. d. Graaff, & P.A.J. Bouhuijs* 1187

Chapter 9
THE SOCIAL CONTEXT OF EDUCATION 1213

Papers
The social context of education, *R.J. Bosker, & A. Wesselingh* 1215
Gender and education: An exploration of research programmes, *E. v. Eck, G. t. Dam, & M. Volman* 1218
Are girls disadvantaged in a computer literate society?, *J. Underwood, S. Cavendish, & G. Underwood* 1221
Determinants of sex - specific subject choices in secondary education, *H. Dekkers* 1226

986

Male and female students at the open university of the Netherlands: Labour-market participation and access to resources, *J. Janssen, & G. Joosten* **1228**
Social inequality and the demand for post-compulsory education: A micro-economic approach, *I. Nicaise* **1231**
Alternative routes from vocational education to the labour market: Labour market effects of full-time vs. dualized vocational education, *R. v.d. Velden, & B. Lodder* **1235**
Vocational schools and the structuring of inequality: A Maltese case study, *R.G. Sultana* **1238**
Social class, gender and vocational training: Qualitative studies of the experience of young women, *I. Bates* **1241**
Students' differential text production in the regulative context of the classroom, *A. M. Morais, & H. Antunes* **1243**
A descriptive study about compulsory school dropouts, *M. Groppo, O. Liverta Sempio, E. Confalonieri, & G. Scaratti* **1245**
Students' gender and social selection in commercial secondary-level education in Finland, *M-L. Stenström* **1248**
A dynamic analysis of the transition from education to employment: The case of Luxembourg, *L-S. Allegrezza-Carvoyeur, & J-L. Kop* **1251**
Young people's developing lifestyles and the transition from school to society: A longitudinal study, *L.B. Hendry, J. Shucksmith, J.G. Love, & A. Glendinning* **1254**
Working practices and childrens' application of subject knowledge in the primary school, *S. Bristow, & C. Desforges* **1257**
Internationalization of education, *J. v.d. Pluijm* **1259**
Learning motivation: intellectual and emotional development of adolescents deprivation, *A.G. Moskovkina, & A.A. Sagdullaev* **1262**
The influence of parents' expectations on school careers, *A.A. v.d. Hoeven-v. Doornum, P. Jungbluth, & M.J.M. Voeten* **1264**
Achievement and achievement motivation, *D. Ardaç* **1268**
The meaning of the occupations of working mothers for the educational inequality, *J. Dronkers* **1271**
Equality and achievement in education after educational expansion, *H. Meulemann* **1273**
Social stratification, political attachment and educational expectations in Hungary: An analysis based on public opinion survey data, *G. Halász* **1275**
Minority language educational rights: A Canadian example, *J.L. Black-Branch* **1276**
The school careers of a third-generation ethnic groups: Molukkan students in the Netherlands, *M.C. Robijns, J. Roeleveld, & A.M. Veen* **1279**
OP-STAP, A home intervention programme preparing children for school: Its effects, *P. Vedder, & L. Eldering* **1282**
OP-STAP, A home intervention programme preparing children for school: Its implementation, *P. Vedder, & L. Eldering* **1283**
Pygmalion and effectiveness of 'black' schools: Teachers' stereotypes and covert target levelling towards minority pupils, *P. Jungbluth* **1284**

987

Differences in policy making capacity and culture: An investigation in two educational priority areas, *I. Berings, H. v. Cuijk, P. v.d. Kley, & A. Wesselingh* 1286

Posters

Achievement in probability of students from opposite sides of the world, *B.R. Zinn, & Y. Voronel* 1289

Earnings, jobs and education in Spain, *R. Martinez, J-G. Mora, & L.Vila* 1293

Symposia

Enhancing educational opportunities, *G. v.d. Werf* 1296

AUTHORS INDEX

988

CHAPTER 7

VOCATIONAL AND PROFESSIONAL EDUCATION AND TRAINING

VOCATIONAL AND PROFESSIONAL EDUCATION AND TRAINING

Wim J. Nijhof, The Netherlands, Frank Achtenhagen, Germany, Martin Mulder, & Jan N. Streumer, The Netherlands

Introduction

This theme emphasises vocational and professional education and training in all its facets. Although there is some overlap, three labour market or job oriented training and education sectors are distinguished broadly:

- initial vocational education;
- corporate training and education;
- continuing (adult) education.

The these is perceived from various (sub)disciplines in educational research, and international comparative, historical, empirical, theoretical and review studies will be presented and discussed. Many issues will be elaborated within the presentations in this theme, such as:

- skill formation and skill shortages;
- determination of job profiles, key competencies, and core skills;
- national vocational qualifications amd curriculum profiles;
- effective relations between the labour market and segments of the education and training systems;
- effects of dualization of vocational education programs;
- learning technology for vocational and industrial education and training, such as simulations, computer-based training for flexible production automation and computer numerically controlled production systems;
- structured on-the-job training and learning processes that go along this training approach;
- transfer of training.

Special events worth mentioning are:
A key note address by prof.dr. Frank Achtenhagen (Georg August Universitat, Gottingen). The title of his presentation is: How should research on vocational and professional education react to the challenges in liofe and in the worksite?
A key not address by dr. David Raffe (University of Edinburgh). His presentation is titled: The "new flexibility" in vocational education.
Four symposia will be organised in which specific problems or topics are examined from a variety of perspectives.

Responsiveness and organisational effectiveness in vocational education and training

Chair: Jan Streumer.

This symposium consists of four papers:

- The development of vocational education and training: industry and service compared (Nieuwenhuis & de Bruijn);
- Responsiveness of vocational education in the printing industry (v.d. Berg & Hövels);
- Developing the quality of responsiveness trough modules (Clark);
- Responsiveness of vocational education towards technological changes: CNC-cutting and Flexible Production Automation (Streumer & Feteris).

Knowledge utilization within vocational education

Chair: Frank Achtenhagen.

This symposium provides four papers:

- Complex teaching-learning environments - theoretical assumptions and practical developments (Achtenhagen);
- Schema-based modeling of complex economic situations (Achtenhagen);
- Knowledge utilization: complex problem solving in commercial education (Mandel, Gruber & Renkl);
- Knowledge utilization: promotion by a training in situated access (Neber).

Issues in developing curriculum programs for corporate training and education

Chair: Martin Mulder.

This symposium consists of four papers:

- From function profiles to proposals for better courses (Versloot);
- A framework for the study of operator competence under technological hazard (Marshall);
- Towards standards in corporate education (Kessels);
- New technologies, shared control and "Union"-education (Feijen).

Discussant: Drs. Martin Mulder.

Education and Training for lower educated adults

Chair: T.F. Brandsma.

This symposium provides five papers:

- Basis Skills and adult education (van der Kamp & Jansen);
- Needs assesment adult education (Doets);
- Transfer from basic education to vocational adult education and adult general education (Moerkamp);
- Company training for employees with low educational qualifications (Onstenk);
- Educational policy and the (non) participation of lower educated adults in adult education (Brandsma).

Besides the key note papers, and symposia several paper sessions will be held: Adult Education, Developments in secondary vocational education, Higher vocational Education, Models for training and development in organizations, Results from human resource development, Apprenticeship systems.

991

CAREER EXPECTATIONS OF STUDENTS OF THE HEALTH SCIENCES

M.L.J.J. Schreurs, Department of Educational Development and Educational Research, University of Limburg, Maastricht, The Netherlands

Introduction

The transition from university to labour market is a difficult period in career development, but is also of vital importance for the students' future careers (Hall, 1976; Schein, 1978). To gain insight into this process of leaving school and entering the labour market, a research project was started. The design is longitudinal. A group of two hundred students are being followed from the last year of their studies until a few years after they will have entered the labour market. The students and alumni have been approached three times during this period of transition from university to labour market: in the last year of their studies, at the start of their first jobs and after the first year in their jobs. These moments of measurement correspond with the first three stages of the process model of the transition cycle (Nicholson and West, 1989): preparation, encounter and adjustment. Central themes of the preparation stage are expectation and anticipation before change. Themes of the encounter stage are affect and sense making during the first days or weeks of job tenure. Central themes of the adjustment stage are subsequent personal and role development to reduce person–job misfits.

The first study in this research project is concerned with the preparation stage. The aim of the study was to gain insight into the expectations of students about their future career and in their orientation to their future career. Van Nieuwenhuysen and Wielers (1990) have discriminated a defensive and offensive orientation to the labour market. Main characteristic of the offensive orientation is a career–directed choice of study. For students with a defensive orientation, choice of studies is not influenced by labour market–figures. Crites (1969) has found that a planning orientation is essential: to work consistently towards one clear goal. In the succes–model described by Hall (1976), an appropriate goal–setting and planning are also core concepts in career development.

To investigate career expectations and orientations of students, we have developed the Career–Expectations Questionnaire. Former study and work experience, future work situation, relation between study and future job, and career planning are the topics of the questionnaire.

992

Later on, the project will investigate whether they have been able to realize their expectations in the encounter and adaptation stage. This paper is confined to the preparation stage of students.

Method

The participants of this study were all the students of the faculty of Health Sciences in the fourth, fifth and sixth year of their studies. The Career–Expectations Questionnaire was presented to 312 students. The questionnaire was filled in and returned by 223 students. There were 56 men and 167 women. Most of the respondents are 23 and 24 years old.

The questionnaire consists of four parts. First, some general features are gathered, like personal characteristics, work experience, major, and duration of study. The second part contains a few questions about their future working life. What is the desired job and which organization is preferred? What do the students expect from their occupation? How do they prepare for their future work situation? The third part is about the relation betweenstudy and future job. In the fourth and last part of the questionnaire, questions about career planning and study satisfaction are asked.

Results

Former study and work experience

A substantial group of the students (18 %) had already studied in other disciplines before they started their studies in Health Sciences. They were for example nurses, social workers, dieticians or teachers. Half of the students had already gained work experience in a relevant field before or during their studies. They worked as research assistants, assistant teachers or trainers, nurses, social workers or mental health workers, as well as in other occupations.

Future work situation

The most desired job was to become an AIO or researcher (31%). The second preferred job was that of a staff member or manager of a department in an organization. Also a considerable group preferred an occupation as a counsellor (15%). Most of the students would like to have their future job in educational and research institutes, hospitals and industrial organizations.

Half of the students would prefer a full–time job. A group of 38% preferred a job of three to four days a week. Approximately one tenth would like to work between two and three days a week. A few students had no idea about the size of their future working week. The students were asked to judge their chances of getting the desired job. Only 44% of the respondents expected to get the job. Half of the respondents answered that they were uncertain of reaching the desired occupation. A group of 8 % did not expect to get the most preferred job. The most prominent reason for this pessimistic view is the situation on the labour market. One student said for example: "The market is limited and there is much competition."

Relation between study and future job

The question about study satisfaction shows that 57% of the students would choose the same study if they were able to choose again. A group of 35 % of the

students were disappointed in their studies and would not again make the same choice if they could choose again. After their studies, they prefer for example psychology, medicine, medical biology and industrial economy.

Career planning

The motive of most students (68%) for choosing to study Health Sciences was the interesting content of the studies. Only 17% of the respondents had the labour market in mind at the moment they chose a particular field of studies. 10% of the students reported university education is a continuation of former study or work experience.

How do students prepare for the labour market during the last year of their studies? Only one quarter of the students had already applied for a job. A group of 28 % of the students intended to write a letter of application in the near future. Nearly half of the students was neither applying for a job at the moment, nor intending to do so in a few months. Reasons for not applying for jobs were for example a desire to prolong the period of study, or to follow another university course. Those who were actively looking for a job wrote four application letters on average. A training in application skills is mentioned by most of the students as very useful (88%).

Students were asked about the importance of different sources of information about future work situations. The most important sources were: people in the field, practicalperiods, contacts and working experience during the practical period. Health sciences teachers were regarded as relatively unimportant as an information source by 40% of the students.

Discussion and Conclusion

A lot of students in Health Sciences had already gained working experience. Nearly one fifth of the students had completed another discipline before they started their studies in Health Sciences. This seems to be a good qualification for starting a career. Besides, most of the students were able to make their expectations concrete. They could indicate the desired job and organization. So one might expect them to be positive about their chances of getting the desired jobs. Unfortunately, this is not what we found in this study. Almost half of the students were doubtful about their future.

One third of the students were disappointed in their studies and would not again make the same study choice. This is a substantial part of the students. One may ask if this percentage is specific for the Health Sciences or if it is a general feature of the preparation stage. May be it is an indication of students' worries concerning their competence to fulfil future role requirements (Nicholson and West, 1989). Subsequent research of the next stages of encounter and adaptation will bring clarity in this issue.

In general preparation for the labour market is very poor. In most cases choice of studies is not influenced by labour market–figures. Only one quarter of the students could be characterized as being offensive in their orientation. They had chosen their studies from the perspective of their future careers. Also one quarter of the respondents had already applied for a job. This is a small group, especially when we realize that they will leave the university soon.

On the other hand nearly all the students recommend the training in application skills. Just before they completed their studies, students felt a lack in skills to enter the working field successfully. Perhaps it is necessary to support students in their orientation towards the labour market. They need information about career development and training in application skills. A course in application skills is presented facultative at the students (Schreurs and Hommes, 1987). But a separate course is not enough. The attitude of all the staff members is important. Teachers should more actively stimulate their students in preparing for the labour market.

References

Crites, J.O. (1969). Vocational psychology. New York: McGraw Hill.

Hall, D.T., (1976). Careers in organizations. Santa Monica: Goodyear Publishing Company Inc.

Nicholson, N., & West, M. (1989). Transitions, work histories, and careers. In: M.B. Arthur, D.T. Hall, B.S. Lawrence (Eds). Handbook of career theory. Cambridge: Cambridge University Press.

Nieuwenhuysen, W.L. van, & Wielers, R.J.J. (1990). De arbeidsmarktpositie van intredende academici. Groningen: RION, Instituut voor Onderwijsonderzoek.

Schein, E.H. (1978). Career dynamics: matching individual and organizational needs. Reading: Addison–Wesley

Schreurs, M.L.J.J., & Hommes, J.A., (1987). Een training solliciteren. In: F.J. Dochy, S.J. van Luyk (Eds). Handboek vaardigheidsonderwijs. Lisse: Swets & Zeitlinger.

WORK EXPERIENCE AND THE DEVELOPMENT OF EXPERTISE

Anneli Eteipelto, Institute for Educational Research, University of Jyvskyl, Jyvskyl, Finland

Introduction

The development of expertise has often been understood as a linear continuum, which has the polarizations found by contrasting the less experieced novices with the more experienced experts. These comparative studies have demonstrated that experts perceive large meaningful patterns of their own domain, and that they see and represent the problem on a deeper level than novices. Experts have also been found to focus quickly on the relevant cues of the task and to have better self–monitoring skills than novices (Chi et al.,1988). The central explanation assigned for these characteristics of the experts' strategies in problem solving are their rich domain knowledge and the qualitatively different organization of their knowledge base.

Most studies of the development of expertise concern the solution of closed problems like physics, which have one best solution, or tasks with testable solutions and little degrees of freedom concerning the aims and goals, like programming. These studies have little relevance or ecological validity for tasks which involve specification of working goals and the choosing of the methods and strategies suitable for the situation. These kinds of tasks are characteristic of the personnel working in planning and developing tasks.

The nature of the expertise of those working with planning and developmental tasks is much discussed and questioned, especially in connection with the development of basic and continuing education. This study focuses on the expertise of those working with information systems design and development. These systems analysts are continously facing new problems of different application fields. Organizational constraints including economic resources and the subjective competence of those taking part in the project constitute the conditions, under which the developmental work is realized.

There are different opinions about the quality of expertise of those working with information systems design. Is the central stress on the methodical expertise, or is it more on the knowledge and deep connoisseurship of the application domain. For the assessment of occupational development and competence it is asked, what it means to develop a comprehensive occupational orientation in the field, and how it is developed through different work experiences. It is also asked how to

996

educate reflective practioners (Sch?n, 1987) who recognize thedevelopmental potentials of work.
All experiences are not utilized as the source of learning. The theory of experiental learning (Kolb, 1984) suggests that experiences get meaningfullness and transform into learning through critical reflection of the experiences or active experiementation with feedback. Dreyfus and Dreyfus (1986) argue that the length of the work experience gets meaningfullness for learning as far as subjects' old conceptions will be tested and challenged in practice.
The development of expertise at the real work settings takes qualitatively different paths. The central area of the development of the expertise depends on the subjects' work orientation, which includes his or her conceptions of 1) the essential topics and issues to be consider in planning and development, 2) the main criteria for the organization of the development process. For the continuous development of expertise, which is an urgent need of the current working life, we need knowledge of the subjects' practical work orientations and their connections with the developmental potentials of the work experiences as the source of learning.
This study which is a part of a larger project analysing the development of expertise of the systems analysts, tries to find and describe the qualitatively different work orientations and their connections with the subjects' work experiences and the utilization of these as the source of learning. The subjects' work orientations are mainly diagnosed from their knowledge structures and from their strategies of solving a planning task.

Method
Information systems students (n=40) and practitioners (N=40) with similar educational background, but different work experiences, were tested and interviewed.
The subjects' work orientation was diagnosed with three different methods. Firstly, the subjects were asked to construct on the blackboard a graphical representation of the information systems development. During the construction process the subjects thought aloud and their construction process was videotaped. Afterwards the model construction process was confronted and discussed with the researcher. Secondly, the subjects were asked to solve a simulated planning task. Thirdly, they answered an essay–form questionnaire.
Subjects were also interviewed about their working and learning histories.

Results
Only preliminary results can be presented at this stage of the study, because the final analyses are not yet completed.
Qualitatively different types of expertise were found among novices and among experts. The most general work orientation among novices was a rather limited and constricted orientation to the the phases of development. Either the methods of development or the end–users were stressed among the novices, but a comprehensive developmental orientation was missing.
Experts with many years of work experience usually had the organizational context and its needs as the starting point of development. Organizational constraints, including economic resources and the competence of those taking part in the

project, were perceived as the concrete conditions for development, and the methods are subordinated to these.
Comprehensive occupational orientation which includes the consideration of the application field, does not always seem to develop in concordance with the increase of work experience.

Conclusions and discussion

It seems that the quality of the subjects' work orientation will crucialy determine, what kind of development takes place when work experience increase.
The results are discussed by comparing the development of the expertise of systems analysts to the development of other occupational groups, like teachers.

References

Chi, M.T.H., Glaser, R., & Farr, M.J. (1988). The nature of expertise. Hillsdale, NJ: Lawrence Erlbaum.

Dreyfus, H., & Dreyfus, S. (1986). Mind over machine. Oxford: Basil Blackwell.

Kolb, D.A. (1984). Experiental learning. Experience as the source of learning and development. Englewood Cliffs, NJ:Prentice–Hall.

Schön, D. (1987). Educating the reflective practioner. London: Jossey–Bass Publishers.

998

FEASIBILITY, RELIABILITY, AND VALIDITY OF USING THE STANDARDIZED/SIMULATED PERSON TECHNIQUE IN PERFORMANCE ASSESSMENT OF COMPETENCE IN THE PROFESSIONS

Nu V. Vu, & Howard S. Barrows, School of Medicine, Department of Medical Education, Southern Illinois University, USA

One main weakness in the assessment of professional skills such as those practiced in the health professions (i.e., medicine, nursing, dentistry), law, or business (i.e. sales, stockbrokerage) etc..., is that these skills have always been measured with written tests. For example, in the area of medicine, written examinations cannot guarantee whether the students will be able to use their knowledge in the care of patients or whether they will perform the needed clinical skills in the appropriate situations. On the other hand, when students' performance is directly observed and rated several problems are identified. First, medical students or residents are rarely observed by faculty (Stillman, Regan, and Swanson, 1987), and when observed there is always a great variability in the patients assigned to the students as well as in the individual faculty members' criteria for rating students' performance. Overall, performance as commonly assessed in most professions until now is subjective and not standardized across students.

It is only after the technique of the simulated/standardized patient (SP) was introduced (Barrows et al.,1964; Barrows,1987) for medicine that the standardization of performance–based assessment with objective scoring becomes possible. The SP is an individual carefully coached by a specific training method to simulate accurately an actual patient. Using the SP technique, faculty are able to standardize both the presentation of the patient's problems to a large number of examinees and the scoring of the examinees' performance. Because the purpose of assessing performance in medicine and in other professions is to determine whether an individual is able to master a series of situations which are common in their practice, a standardized patient can be easily translated into a standardized client or board of directors for example, to assess performance in the professions in an objective and standardized manner. The purpose of this paper is to report seven years of results obtained on a clinical practice examination (CPX) using standardized patients. The paper specifically reports on the feasibility, reliability and validity of the CPX.

Method

Subjects and Instrument

A clinical practice examination or CPX is administered in the last seven years (1986 – 1992) to fourth–year medical students (N=475) at the end of their clinical clerkship rotation to determine their overall clinical competence and readiness for the residency training.

Each CPX has between 13 to 18 SP–based patient problems, with the problems and the skills assessed within the problems sampled from the students` Exit Objectives. The Objectives list all the skills, as well as the presenting complaints and diagnoses which a student should be able to evaluate and/or manage by graduation. For each problem, the examinees have a 20–minute clinical encounter with the SP. During the encounter, the examinees are assessed on their abilities to perform an appropriate history and accurate physical examination, to educate the patient, and/or to interact and relate with the patient. In addition to simulate the patient problem, the SP is trained to record on a detailedbehavioral checklist those behaviors which the examinees did or did not perform. After encountering the patient, the examinees are given a 20–minute written or computerized test aimed at assessing what in real practice is referred to as the physicians' case follow–up. This consists of assessing examinees' skills in eliciting patient data, evaluating the patient's problem, ordering relevant laboratory and diagnostic tests, interpreting the test results, and/or managing the patient.

Each of the patient case has a scoring key and a passing score set by the case's clinical author, which are then reviewed and adjusted, if needed, by a committee of physicians of different clinical disciplines. The students' performance on the CPX is reported by their average case score and whether they pass or fail the exam. A student passes or fails the CPX if s/he average case score is respectively higher or lower than the average case passing score.

Results

Overall, the CPX administration is not only feasible but also transportable in that it can be administered at a different site with reasonably high accuracy and standardization in the simulations (Klass, 1987; Tamblyn et al. 1991a). The mean case scores (range = 43 to 91) and standard deviations (range= 5 to 19) of the CPXs indicated that the examinations capture a wide range of performance. Regarding the CPX test security, it was found that when the exam is administered to five groups of students over a period of three weeks, there was no consistent pattern of differences between the overall as well as the case–by–case mean performance in the five groups (Colliver et al., 1991).

Because with this examination, the accuracy and reliability of the SP's simulation and checklist recording of performance were crucial to the overall test validity and reliability, several studies were performed. By having the SPs' simulations and checklist recordings reviewed and verified by independent observers, it was found than when the SPs were carefully trained, they were highly accurate in their simulations (92%) (Tamblyn, 1991a), and that the use of multiple SPs to simulate the same case did not affect the students' total test scores reliability or reproducibility. It was also found that the SPs' checklist recordings were reliable and accurate (80%) and that their accuracy did not change over the course of a one–day or three–week examination (Tamblyn et al. 1991b; Vu et al., 1992). When

1000

the SPs are used to rate the examinees' communications skills and professional behavior and service, their ratings were found to be highly generalizable (.69 to .83) when based on 15 to 18 patients (Vu et al.,1990).
In view of whether the use of SPs introduces any type of test biais it was found that the the SPs' gender and age did not have a main effect nor interact with the examinees' gender to affect their test performance.
Based on 13 to 18 patient cases, the reliability (generalizability coefficient) of the overall test score was moderate; it ranged from .56 to .72 with an average of .64. The moderate reliability is typical of the reliability found with other formats of assessment of clinical competence or professional competence. The moderate reliability is often attributed to what is referred to as case specificity, that is students' performance varies with the patient cases. Because the CPXs are used to make pass–fail decisions, it is more relevant to determine the generalizability of the pass–fail decisions with cut–off score which ranged from .55 to .93 for the seven CPXs or an average of .84.
With careful selection and development of the patient cases it was found that the examinations had content validity in that they assessed in average 84% of the skills listed inthe Exit Objectives. It also had content fairness in that for each of the CPX case a small (13%–37%) to a large percentage (71%–89%) of students indicated that they have been directly exposed or worked up before a similar problem, and a moderate (56%–75%) to a high percentage (98%) indicated that they have read about it. The examinations also demonstrated construct validity in that examinees at different levels of training performed differently on the examination (Klass et al., 1990; Barnhart et al., 1992). In addition, students had indicated on a questionnaire survey that the CPX assesses their ability to use their clinical and problem–solving skills (Shirar et. al, 1992). Regarding the utility of the CPXs in predicting students' performance in their first–year of residency, it was found that the CPXs are sensitive but not as specific (Vu et al., 1992). These results are still preliminary and since a large percentage of residency is missing data for those students failing the CPXs.

Conclusions

Overall the results of performance assessments using the simulated/standardized patients to assess competence in medicine indicate that this type of examination is feasible and demonstrates reasonable levels of reliability and validity. The results are encouraging for the use of the SP technique in any professions where the candidates need to be assessed on their abilities to work with different clients and master different practice situations.

References

Barnhart, A. et al. (1992). A comparison of second–and fourth–year medical students on a standardized–patient (SP) examination of clinical competence: A construct validity study. Presentation at the American Educational Research Association Annual Meeting, San Francisco.

Barrows, H.S. (1987.). Simulated (standardized) patients and other human simulations: A comprehensive guide to their training and use in teaching and evaluation. Chapel Hill, NC: Health Sciences Consortium.

Barrows, H.S., & Abrahamson, S. (1964.). The programmed patient: A technique for appraising student performance in clinical neurology. Journal of Medical Education, 39:802–805.

Colliver, J.C. et al. (1991). Test security in examinations that use standardized patient cases at one medical school. Academic Medicine,66:279–282.

Klass, D. et al. (1990). Influence of level of training on performance in a standardized test of clinical abilities. In Bender, W. et al. (Eds.). Teaching and assessing clinical competence, The Netherlands: Boekwerk publishers.

Klass, D.et al. (1987.). Portability of a multiple station, performance based assessment of clinical competence. In I.R. Hart, and R.M. Harden (Eds.), Further developments in assessing clinical competence. Montreal, Canada: Can–Heal.

Tamblyn, R.M. et al. (1991a). The accuracy of standardized patient presentation. Medical Education, 25:100–109.

Tamblyn, R.M. et al. (1991b). Sources of unreliability and biais in standardized patient ratings. Teaching and Learning in Medicine,3:74–85.

Vu, N.V. et al. (1992). Checklist characteristics and length of testing: their effects on standardized patients' simulations. Journal of Medical Education, 26, (In press).

Vu, N.V. et al. (1992). Performance–based test sensitivity and specificity in predicting first–year residency performance. Presentation at the American Educational Research Association Annual Meeting, San Francisco.

Vu, N.V. et al. (1990). Generalizability of standardized patients' satisfaction ratings of clinical encounter with fourth–year medical students. Academic Medicine, 65, S29–S30.

1002

MODULAR INSTRUCTION, STUDENT INVESTMENT AND SCHOOL OUTPUT

G.J. Meesterberends-Harms, Institute for Educational Research (RION), Groningen, The Netherlands

Introduction

During the past years modules have gained importance in Dutch vocational education. The Department of Education propagates modular instruction in a statement published in 1988. Two of the ten goals stressed are:

- to raise output of courses and
- to improve students motivation.

The research study to be reported here was carried out in order to examine whether (different types of) modular systems have the supposed effect on output and students' motivation. Hypotheses are:

- Under certain conditions a modular system induces a higher student investment in school. These conditions are to be found in: (1) characteristics of the modular system itself; (2) student characteristics; (3) characteristics of the class or instructional group to which the student belongs.
- Student investment in school affects course output;
- Modular systems increase course output; this increase can mainly be explained by an increased amount of students investment.

Two types of vocational education were examined: full time short vocational education (Kort MBO) and the part time school component of the apprentice system (BBO).

Methods

There are three treatment groups:

- pupils in a non-flexible modular system (variant 1);
- pupils in a flexible modular system (variant 2);
- pupils in a traditional, non-modular system (variant 0, control group).

The total number of students is 776.

The following variables were measured:

(Modular) variant (school level, independent variable), achievement motivation, level of cognitive functioning, and experienced motivating effect of subject content (all student level covariates), task supporting character of the instructional group (school level, covariate), student investment (student level, dependent variable in some of the analyses, independent in others), output (student level, dependent variable).

The analyses were based on modelling the supposed effects by linear hierarchical models. In order to do this the multilevel techniques of the VARCL-program were used.

Results

Students in the apprentice system, who work in a modular variant, show a higher investment-score after three to four months of education than students in the same system who work in a non-modular variant. The flexibility of the system does not affect investment at this point of time. In the long run, that is over a one year period, the non-flexible modular variant works best for student investment: the flexible variant does not distinguish itself from the non-modular variant in this respect.
Students in full-time short vocational education who work in a flexible modular variant show a higher investment-score after three to four months of education than students in the same system who work in a non-flexible modular variant. The same difference between a flexible and a non-flexible modular variant remains when analyzed over a whole year.
The reported effects exist after having controlled for the influence of the covariates.
In any of the three moments throughout the year in which it was measured, student investment has a positive relation with student progress as measured after one year of education. Student investment as measured after three to four months of education has a negative relation with drop-out within one year. This holds for both types of education under ivestigation.
The (modular) variant never has its own additional effect on any of the two output measures, once the influence of student investment is controlled for. There is one exception: in short vocational education, after having controlled for student investment, students who work in the flexible modular variant show less progress than students who work in the non-flexible modular variant.

Conclusions and discussion

The results in general support the first hypothesis: modular instruction leads to more student investment. To produce this effect there seems to be hardly any need for fulfilling further conditions. Only in one of the two types of education there is an additional positive effect of modular instruction if it the system is more flexible. Characteristics of pupils or instructional group do not matter.
The differential effect of the flexible modular variant for the two types of education is surprising. Causes are sought in the difference in amount of educational contact time or in possible differences in pupil characteristics not accounted for, like age or successful completion of an earlier stage of education.
The second hypothesis is fully supported by the results. In both types of education student investment leads to less drop-out and more progress.
The third hypothesis is for the main part supported by the results. Amelioration of output takes place by increasing student investment; in most cases there is no additional influence of the (modular) variant, except for the influence on progress in full-time short vocational education.
It seems that in this case the flexibility of the system allows students on the whole to take more time for their study than they can get in a less flexible system. It might be useful to examine whether this extension of study time is productive and for that reason necessary or not.

1004

KEY QUALIFICATIONS AND INTERMEDIATE VOCATIONAL EDUCATION (MBO)

Simone J. van Zolingen, & B. van Onna, University of Nijmegen, Nijmegen, The Netherlands

The transition from school to work has come under pressure during the last ten years. New technologies, changes in the structure of organizations, and in the labour market, and ruthless competition affect the qualifications employers ask from school-leavers. Qualification demands are also changing. Nowadays employers need a well-trained, flexible workforce, that is capable of adapting itself to continuous changes. Employers often find school-leavers lacking in knowledge of new technologies, in personal skills and understanding of working life .

The problem we adress in this contribution is the above mentioned discrepancy between the qualifications needed by the world of labour and the qualifications acquired during intermediate vocational education. For the schools the difficulty is, that different companies ask different qualifications from school-leavers. However, intermediate vocational education is obliged to teach qualifications, that are not to specialized, so that school-leavers will be able to work in companies of various organizational set-ups.

Mertens (1974) proposed the concept of *key qualifications* as a solution for this problem. According to him intermediate vocational education should concentrate on teaching broad and basic qualifications, such as logical thinking, the ability to handle information efficiently, methods of self controlled learning, basic technical and social knowledge and so on. The aim of intermediate vocational education becomes the training of flexible school-leavers, who can quickly specialize and adapt themselves by training on or off the job.

In our research project we adress this problem. Our project concerns on the one hand school-leavers that have finished intermediate technical schools with a specialization in electronics (MTS-E) and now work as service mechanics, and on the other hand school-leavers that have finished intermediate business education with a specialization in commerce (MEAO-C) and now work in banks as bank clerks or in insurance companies as claims assessors or as claims acceptors. We have confined this study to two educational specializations and to three occupations, to enable us to relate the contents of 24 selected key qualifications to the selected occupations .

We describe key qualifications as *those qualifications (knowledge, understanding and skills) that belong to the core of an occupation or group of functions and can be transfered to new functions within the occupation or the group of functions.*

We selected 24 key qualifications using information from interviews with a number of employers, interviews with school-leavers from intermediate education holding entrance jobs, and information from other research projects: (1) thinking in the abstract, (2) methodical thinking, (3) problem solving, (4) planning of work, (5) handling information, (6) knowledge of subjects outside the core knowledge of one's profession, (7) self-reliance, (8) sense of responsibility, (9) precision, (10) decisiveness, (11) taking the initiative, (12) handling stress, (13) creative thinking, (14) manual instrumentation skills (only MTS) (15) complying with security measures, (16) workmanship, (17) feeling for commerce and marketing (only MEAO), (18) communication skills, (19) expressing oneself in writing, (20) expressing oneself orally (only MEAO-C), (21) knowledge of modern languages, (22) a positive approach to continuing (self)education, (23) representing a company,, (24) knowledge of the organzation, (25) a critical attitude.
Our research project, this contribution is based on, contains three phases:

- A pilot study among 18 school-leavers of intermediate vocational education, specialized in electronics or commerce, holding entrance jobs (till 5 year after leaving school). For this pilot study the school-leavers were extensively interviewd (2 to 3 hours). They were asked to describe the contents of their jobs and relate this description to 24 specifically mentioned key qualifications, in order to arrive at a description of each of the 24 key qualifications. They were further asked to indicate the importance for their jobs of the 24 key qualifications. Also they were asked when, in their view, these 24 key qualifications can best be obtained. Further 11 job experts and 8 educational experts were extensively interviewed about the same questions and about the influence of changes in organizations, changes in the labour market, and the use of new techniques on the jobs of the school-leavers and on the curriculum of the schools. These 37 interviews were preceded by a literature study about 'key qualifications'.
The material from this pilot study was used as an information basis for the Delphi technique, that takes up fase 2 and 3 of the research project:
- A questionnaire send to 38 experts (among them 10 educational experts with knowledge from intermediate technical schools, 10 educational experts with knowledge from intermediate business education, 14 job experts from industrial organizations and 14 job experts from banks and insurance companies). The experts are asked to fill in the questionnaire in order to give their opinion on the research questions .
- Two workshops. The workshops will follow the filling in of the questionnaire. There is one workshop for all (job experts and educational experts) experts of the industrial sector and one workshop for all (job experts and educational experts) experts of the service sector.

In this contribution we present a few results of phase 1 and 2 of the research project. In the first place results concerning the question:

Which of 24 specially mentioned key qualifications, are important for school-leavers from intermediate vocational education with a specialisation in electronics or a specialisation in commerce, holding entrance jobs?

At this moment only the results of the school-leavers and job-experts are available. At the conference the results of the educational experts will also be presented and compared with the results of the other groups. The respondents were asked to give

1006

their opinion on the importance of key qualifications on a 5 point Likert scale. For every key qualification the mean for each group (school-leavers, job experts) was calculated and key qulifications with mean>4,5 were labeled most important.
School-leavers of intermediate technical schools (N=10) mention as the most important key qualifications for their entrance jobs: self-reliance, sense of responsibility , thinking in the abstract, precision, problem solving, and decisiveness.
School-leavers of intermediate business education (N=8) mention as the most important key qualifications for their entrance jobs: handling information, expressing oneself orally, self-reliance, precision, communication skills, problem solving and workmanship.
The job experts of industrial companies, banks and insurance companies (N=32) all mention as the most important key qualifications for school-leavers of intermediate vocational education: a sense of responsibility self-reliance, and communication skills. These results are consistent with the results of De Jong et al. (1990). They also report that job experts value a sense of reponsibility and communication skills higly.
The job experts of industrial companies (N=14) specifically mention: problem solving, precision, a positive approach to continuing (self)education , methodical thinking, taking the initiave, complying with security measures, workmanship, decisiveness, creative thinking, communication skills, and handling information.
The job experts of banks (N=14) specifically mention: a positive approach to continuing (self)education, workmanship, representing a company, precision, complying with security measures, feeling for commerce and marketing, handling stress and expressing oneself orallly.
The job experts of insurance companies (N=4) specifically mention: decisiveness, taking the initiative, expressing oneself in writing, and expressing oneself orally.
The opinions of the school-leavers and the job experts of the industrial sector do not show much difference. But the job experts mention more important key qualifications. The opinions of the school-leavers and the job experts of the service sector show more discrepancy. The reason for this difference between the two sectors may be that occupations in industry are easier to describe specifically than occupations in the services.
The second question we adress in this contribution is:

Where should these key qualifications be obtained and according to whom?
The second question concerns two different approaches to the transition from school to work, the direct linking approach and the flexible linking approach respectively.
The underlying assumption of *the direct linking aproach* is, that intermediate vocational education should base its curricula on the demands expressed by the employers. To this end representatives of the field of education and representatives of all branches of business confer on national level. On the basis of information about work, occupations and functions given by representatives of business, occupational profiles are developed. These occupational profiles are subsequently translated into educational profiles and eventually in final attainment levels which will be affirmed by the Department of Education. Curricula

based on these final attainment levels, will then be developed. These can be used and elaborated upon by schools in intermediate vocational education.
The underlying assumption of *the flexible linking approach* is, that qualifications are learnt both at school ànd at the workplace. The learning of key qualifications is part of alife long learning process.Companies should also bare responsibility for the transition from school to work and life long learning. This responsibility can be expressed, for instance, in creation of the possibility of learning on the job by preventing the division of labour, and the creation of the possibility of following internal and external training courses . Another assumption is, that it is harder and it takes much more time to adapt whole school curricula to quick changes in organizations and technologies, than to adapt short training courses or the workplace itself.
So the second question concerns the opinion of the job experts and the educational experts on where key qualifications should be obtained. Do they think about the transition from school to work from the viewpoint of the direct linking approach, or from the viewpoint of the flexible linking approach?
The results related to the second research question where inconsistent.
When generally asked, the job experts called themselves proponents of the flexible linking approach. But, when filling in a question about where the 24 different key qualifications should specifically be obtained, the job experts indicate that all of the key qualifications, they mentioned as most important for shool-leavers of intermediate voacational education, should in the first place be obtained at school or cannot be obtained at all. Only decisiveness should according to job experts of banks and insurance companies in the first place be obtained after school, and handling stress and feeling for commerce and marketing should also in the first place be obtained after school according to the experts of insurance companies.
Just like De Jong et al. (1990) our results indicate that key qualifications such as expressing oneself in writing, precision, handling information and something less communication skills should be obtained at school or at a traineeship.
From these results it can be concluded that the direct linking approach and the flexible linking approach do not exclude each other, but exist together.

References

Jong, M.W. de, Moerkamp, T., Onstenk, J.H.A.M., & Babeliowsky, M. (1990). Breed toepasbare beroepskwalificaties in leerplan en beroepsprakrijk. SCO: Amsterdam.

Mertens, D. (1974). Schlüsselqualifikationen. Thesen zur Schulung einer modernen Gesellschaft. Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, Z, 314-325.

WORK SIMULATIONS IN DUTCH ADULT VOCATIONAL EDUCATION CENTRES

Karen Brouwer, Center for Applied Research on Education (OCTO), Department of Education, University of Twente, Enschede, The Netherlands

Introduction

In recent years interest in work simulations as a part of vocational training has increased. A work simulation, e.g. an office simulation, is a simulated administration section of a business office. In such a simulation a learner can practice all relevant skills (of which the performance level depends on his or her alleged capacities) that are necessary for working in a real administration of a business office. Work simulations are specially usefull for training of practical skills as social skills, problem solving skills and attitudes (Alessi & Trollip, 1985).
Treep and Pieters (1989) made an inventory of work simulations in the Dutch regular vocational education and in the adult education, as a result of which the Dutch Organisation for the Provision of Labour (Arbeidsvoorziening) asked for an inventory of the work simulations within their administration trainings. In the Adult Vocational Administrative Training Centres (Centra voor Vakopleiding voor Volwassenen), which provide trainings for unemployed people who have to be (re)trained, work simulations are a regular part of the trainings.
In the eighties, the centres individually developed various types of work simulations. As a result, many different models underlie these simulations. One of the most obvious distinguishing characteristics is the simulation of the external relations with the outer world. On the one hand, external relations can be simulated by a script or by an "order bank"; on the other hand, the outer world includes other work simulations. The first type of work simulation (*the closed–loop system*) restricts the learner in behaving, the latter type (the open–loop system) allows the learner to take initiative to interactively deal with other customers. There are also Centres with a mixed system, which is an attempt to combine the advantages of both models.
Though the representation of reality is mostly the best worked out aspect of work simulations, there are other important distinguishing features. To systemize the results of the inventory a theoretical framework is deduced from studies in computer simulation. De Jong (1990) has described four attributes by which (computer)simulations can be characterized:

1009

- Presence of formalized, underlying models;
- Presence of instructional goals;
- Elicitation of specific learning processes;
- Presence of learner activity.

This presentation concerns three questions:

- What are the main characteristics of the work simulations used in the Centres: a.o. the underlying model, the instructional goals, the elicitation of specific learning processes, and the presence of learner activity, and also the amount of reality and the embedding of the work simulation in the curriculum of the Centres?
- In what way do the Centres succeed in the integration of theory and practice, being a main educational goal of the Centres?
- How do the Centres with an open–loop system organize the cooperation in trade and what is the role of the simulation supervisors, whether their role being on–line supportive or off–line?

Method

Data have been collected in 20 administrative Adult Vocational Training Centres, two technical Adult Vocational Training Centres with an administrative department and two Centres for Vocational Orientation and Training Courses (C.B.B.). The work simulations of these centres have been joined by contacts between supervisors as well as by simulated trade.

In the first phase, a written questionnaire about the general characteristics of the centre and the trainings has been sent to the training head in each centre who is responsible for the educational organisation of the curriculum. By means of structured (closed) questions information of the courses, the organisation and other statistics of the centre has been gathered.

Secondly, one (sometimes two) of the work simulation supervisors of each centre has been interviewed about the characteristics of the simulation. Simultaneously, a within–site analysis has been carried out in each work simulation, including a description of structure, the physical characteristics, and the functioning of students. The interviews are based on a structured interview scheme, which contains the following divisions: organization of the work simulation; differences between work simulations within a centre; external contacts between work simulations; working procedure; supervision; evaluation; structure and furnishing of the work simulation and used instructional materials.

The quantitative data from the verbal reports have been analysed by means of descriptive statistics (medians, percentages etc). The qualitative data have been clustered by means of a qualitative technique and the relations are analysed.

Next, for implementational reasons the results have been presented at two meetings; one with the managers of the centres and one with the training heads and the work simulation supervisors. Results have been discussed and information has been added to the existing qualitative data–base. Also, during these meetings plans have been made for future cooperation between the centres.

1010

Results

Model

The model of most work simulations is an exact copy of a real office department (mostly of a wholesale house). The amount or level of reality ought to be as high as possible. Most work simulations have been developed based on the practical experience (in jobs) of the supervisor. As a consequence, most attention is payed to the representation of a company, not to the learning activities. The educational models of the work simulations are therefore fuzzy and ill–structured. The premisse is that students who are working like reality, are learning. Supervisors consider a high amount of reality as the very important learning experience. For work simulations, the organisation models are almost always a combination of a closed and open loop model. Most work simulations use an external office to simulate the outside world, like customers and suppliers, a bank, an assurance company, etc. The equipment of the simulation offices is modern and according to the supervisors almost identical to that of real companies.

Instructional objectives

There are almost no explicitely stated objectives or instructional goals, written down in official documents. The tasks are analyzed and these are described in terms of assignments for the students. Implicitely stated: to feel how working–life is, to cooperate with colleagues, to plan, to adequately deal with apparatus, to carry out assignments, to manage work.

Learning processes

Instruction and learning processes in work simulation include demonstration or modelling by the supervisor, a gradually decreasing coaching (fading), acting with confrontations and feedback, control by peers.

Learner activity

Learning activity is implied in open loop systems more diverse and varied than in closed loop systems. In closed loop systems the students are sometimes allowed to deal with others. In open loop systems students act with much more responsibility and initiative.

Integration of theory and practice

The integration of theory and practice, although stressed by many training heads and supervisors, is often a problem. There is few discussion between work simulation supervisors and theory teachers. They often do not know what is going in each other lessons.

Cooperation between work simulations and between supervisors

There are several attempts to regulate the simulated trade between the open–loop simulations. Examples are the formation of central services as a bank and an assurance company. All work simulations can join this services. Many supervisors expressed the need for more cooperation between the work simulations, both on trade and the cooperation between supervisors.

Discussion

There are many kinds of work simulations, organized on various models. Because of the lack of expressed instructional goals, methods and criteria, it is difficult to compare the work simulations on these aspects. Therefore we have developed a descriptive tool, based on the theoretical framework. With the help of this tool, the supervisors and training heads can make a systematic description of the work simulation and it's place within the curriculum.

Another item in future research is the development of assessment tools with which especially the practical skills can be assessed.

References

Alessi, S.M., & Trollip, S.R. (1985). Computer–based instruction: Methods and development. Englewood Cliffs: Prentice Hall.

De Jong, T. (1991). Instructional environments for simulations. Education & Computing, 1, 305–358.

Treep, M.J., & Pieters, J.M. (1989). De praktijk leren door simuleren. Enschede: Universiteit Twente, Onderzoekcentrum Toegepaste Onderwijskunde.

1012

EFFECTS OF ADULT EDUCATION: OVERVIEW AND PERSPECTIVES OF RESEARCH IN THE NETHERLANDS

J. Boon, & A. van Staa, Department for Research and Evaluation, Open University, Heerlen, The Netherlands

Introduction

Until recently research into the field of adult education was principally centred on participant characteristics (such as gender, age, labour-market position, educational level), internal efficiency of the institutions and drop-out figures of participants (Houtkoop, 1985). The lack of ex-post evaluation of effects of adult education for students can be explained by several factors, as the small allocation of funds, the large dispersion in training and educational programmes and the heterogeneity of participants often not seeking certification. Besides these factors the absence of ex-post evaluation of effects can also be related to the tempestuous development this educational field has experienced the last 20 years, which was accompanied with changing policy-objectives.

In this paper we will give an overview of research into the effects of adult education in the Netherlands and we give a short description of our own research in this respect.

In our view research into effects should be more theoretically anchored.

A theory about effects of adult education should not only take into account costs and expectations of individual actors but should incorporate the importance of the (educational)environment, of relations and networks in decision making on participation in adult education.

Also we argue that linkage of questions on participation and on effects is necessary. The hypothesis can be formulated that the main effect of adult education is the change or the stability in the relative position of a participant in his network or environment in those situations where education and training is valued as a vehicle for mobility.

For research into training effects a consequence of this theoretical model is that attention should be given to the function and the value education has in the environment of participants; f.e. educational level of colleagues and members of the network of participants.

Overview of research of adult education in the Netherlands

In the seventies non-formal and out-of-school programmes dominated the field; two other forms of adult education, formal education and on-the-job training, developed separately. Research into adult education was mainly focused on direct

application in the field, and was often carried out by order of decision-makers as a preparation of the development of new programmes. Most research was explorative and descriptive.
Throughout the eighties the field of adult education changed profoundly. The division into different forms of adult education was generally seen as unsatisfactory and a legislation procedure was started to harmonize the system. At the same time adult education was no longer mainly seen in cultural-political terms of emancipation for disadvantaged groups, but also training for unemployed and, more generally, labour – market objectives became important. Since the mid-eighties, we can observe a change in research into adult education and a growing importance of vocational programmes and on-the-job training. The nineties will probably show a further growing together of different forms of adult education, and a blurring of the boundaries between adult education and vocational training.
This new interest influenced adult educational research. The research field was broadened and more attention was given to economic cost benefit analysis. Up to now researchers do not agree on the definition, the measurement and even the measurability of effects.
Two central themes are covered in Dutch research.
First we find research into effects on mobility and social inequality.
The main stream of research though is on the labour-market effects of different certificates or kinds of training. Looking at the theoretical background of these studies a strong attention for the human capital approach, measuring effects in terms of income be seen.
Although a great deal of the supply of educational facilities is inspired by a policy aimed at diminishing educational and social inequality, few research projects are set up to the evaluation of these effects.
Using longitudinal data, Van Leeuwen and Dronkers (1991) find a non-linear relationship between youth education and participation in adult education. Men who only attended primary school and men with a university degree participate less.
Analyzing another longitudinal dataset, Diederen (1991) describes the careers of a cohort between the age of 25 and 35. 83% of the cohort participated in one or more courses after completing youth education. Effects can be seen on the job level and the enterprise sector. This effect although is mediated by other characteristics of the student as previous educational level and sexe.
The second field of research is about formal enterprise training and on-the-job training. Last years we can see a growing interest in this field. Income effects of on–the–job training are compared with effects of formal training (Groot, Hartog, Oosterbeek, 1991).
It is clear that not all groups of workers have equal chances to participate in training. A positive bias can be seen for characteristics as employment time pro week, sex, years of schooling and job level (Teulings, 1990).
Participants themselves describe the effects of training often in terms of better functioning in the job or a growth in selfconsciousness, and in better position on the internal and the external labour-market (Wong, 1987).
A specific kind of training consists of pre-employment training. Education and courses positively affect reentry chances for women, immigrants and younger

1014

employees. No effects can be found for reentry of unemployed aged 35 or older (Ridder, 1986).
At the beginning of this overview we formulated a hypothesis on the relevance of networks and the educational environment in the decision making and the effects of participation. The defence of ones relative position can be seen as an important factor in the decision making about investing in education as an adult. We tested this hypothesis in a research on students of the Dutch Open University. A group of students is compared to a group of non-students.
Results point to the acceptance of our hypothesis; the educational level of members of the environment and the extent of ones network are important determinants for participation. The effects of study can be interpreted accordingly, students mention relational effects as better functioning, more confidence. (Boon, 1990).

STRUCTURED ON-THE-JOB TRAINING IN THE NETHERLANDS

J.A. DeJong, A.S.A. Vendrig, & A.M. Versloot, Department of Education, Utrecht University, Utrecht, The Netherlands

The importance of on-site training is clearly illustrated by the degree to which it is used. On-the-job training (OJT) constitutes an important part of the total training effort of companies (according to Sloman, 1989 at least 50 per cent). Much OJT takes place in an informal manner; new personnel is incidently coached by their supervisor or colleagues (unstructured OJT). Yet some companies have developed more structured forms of OJT and some research has been done to evaluate these forms of training (Belbin, Belbin and Hill ,1957; Cullen et al., 1976, DeJong, VanElst & Pool, 1991, Jacobs, 1991, Jacobs and McGiffin, 1987, Jansen, 1990, Kolkhuis-Tanke, 1990). Structured on-the-job training has been defined as "*...a task-oriented training (or part of such a training) that takes place on the shop-floor. The learner carries out assignments according to a written plan, being coached by an experienced colleague or by a supervisor. At the end the mastery of the tasks is formally assessed*".

Although Mulder et al. (1988) pay attention to OJT in their survey of practices in Dutch corporate education, little is known of the extent to which structured OJT is being practiced in Dutch firms.

The study presented in this paper is part of a larger research program. We give a sketch here of the theoretical embedding of the study, and an outline of previous and forthcoming studies within the program.

This paper presents the results of telephone interviews held with training and development managers of the 149 largest commercial firms in the Netherlands. Questions were asked about the occurrence of structured on-the-job training, plans to develop such training programs, the percentage of workers that received structured on-the-job training, and the way on-the-job training related to off-the-job courses. The training and development managers were asked to send descriptions of their programs and materials used. Thus, an impression was formed of the training programs as they were used in the firms.

A list containing 211 names and addresses of firms with more than 750 employees, was provided by the Dutch Chamber of Commerce. This could be reduced to 149 firms by excluding duplicate entries and small firms erroneously included in the list. Most of the companies listed were industries, and banks and insurance companies.

The data presented below are provisional. Final data will be presented at the conference.

1016

Structured on-the-job training, as defined, is practiced slightly more often in industries (54 %) than in banks and insurance companies (40%).
In the subgroup of firms that already offer structured OJT, three in four intend to extend their structured OJT programs. One in four firms that offer no structured OJT intend to develop such training within this period.
Structured on-the-job training is usually offered to a relatively small proportion of staff only: one in two firms that offer structured on-the-job training do so for only 20% of the work force or less. Most structured on-the-job training was offered to blue collar personnel in technical and/or production jobs.
Structured on-the-job training can be linked with training off the job in three ways:

- closely related;
- loosely coupled;
- without additional off-the-job training.

The paper gives data concerning the frequency of occurrence for each type of linkage.
The paper concludes with some theoretical remarks concerning the appropriateness of structured on-the-job training in corporate education, and the relation of such form of training to human resource management.

References

Sloman, M. (1989). On-the-job training; a costly poor relation. Personnel Management, February 1989, 38-41.

Belbin, E., Belbin, R. M., & Hill, F. (1957). A comparison between the results of three different methods of operator training. Ergonomics, 1 (1), 39-50.

Cullen, G., Sawzin, S., Sisson, G. R. & Swanson, R. A. (1976), Training, what's it worth?; an experimental case study at Johns-Manville Corp. Training and Development Journal, 30 (8) 12-20.

DeJong, J. A. de, VanElst, G., & Pool, I. (1991). Werkplekopleiden bij Hoogovens IJmuiden. Utrecht: Rijksuninversiteit Utrecht; Vakgroep Onderwijskunde.

Jacobs, R. L. (1991). Forecasting the financial benefits of unstructured and structured OJT. Paper presented at the International Conference on Research on Corporeate Training and Development at Enschede (Nl.), Sept. 27, 1991.

Jacobs, R .L., & McGiffin, T. D. (1987). A human performance system using a structured on-the-job training approach. Performance and Instruction, juli 1987, 8-11.

Jansen, M. (1990). Evaluatie opleidingen monteur en hoofdmonteur seinwezen; werkplekopleidingen getoetst aan de praktijk. Utrecht: Rijksuniversiteit Utrecht; Vakgroep Onderwijskunde (Master thesis).

Kolkhuis-Tanke, I. R. (1990). Evaluatie proefopzet: flexibele opleidingsorganisatie Lijnwerkplaats Amsterdam (unpublished report). Utrecht: Nederlandse Spoorwegen.

Mulder, M., Akkermans, J. S., & Bentvelsen, N. (1988). Bedrijfsopleidingen in Nederland. 's Gravenhage: Instituut voor Onderzoek van het Onderwijs.

1017

QUALIFICATION-POLICY IN PROFESSIONAL ORGANISATIONS: AN EMPIRICAL EXPLORATION OF SOME THEORETICAL PROPOSITIONS ABOUT LEARNING - AND QUALIFICATION -STRATEGIES IN DIFFERENT TYPES OF ORGANISATIONS

Ferd J. van der Krogt, Vakgroep Algemene Pedagogiek, & John Warmerdam, Instituut voor Toegepaste Sociale Wetenschappen, Katholieke Universiteit Nijmegen, Nijmegen, The Netherlands

Research on education and training in labour organisations shows that there are many differences between different types of organisations in the way they structure, i.e. give form and content to, the processes of acquisition and development of qualification of their employees. However, discussion on the question of how to explain these differences has rather been scarce. In this paper we try to tackle this question from an organisation-theoretical point of view by making a connection between the organisation-theory of Henry Mintzberg and qualification-policies in labour organisations. Starting with Mintzbergs' theory we develop some hypotheses on the relations between configurations of organisation and qualification-policy. For one type of organisation, the professional organisation, we make an assessment as to what extent these hypotheses are tenable. This assesment is based on an empirical investigation conducted in four professional social service institutions. For each of these institutions we describe the organisation of the learning function and the learning programmes. We also compare the four institutions on these points. The similarities and differences which are observed are then discussed in relation to the earlier formulated hypotheses.

Theoretical backgrounds

When there occur discrepancies between required qualifications (i.e. qualifications needed to do the work within an organisation) and qualifications at hand (i.e. qualifications at the disposal of present or new employees), in principle there are three possible ways to reduce them:

a strategy of adapting qualifications at hand to the required qualifications. In this strategy qualifications of employees are adapted to (new) tasks and functions by education and training;

a strategy of adapting required qualifications to the qualifications at hand. In this strategy the qualifications of employees, gained by former training and

1018

workexperience, are taken for granted and tasks and functions are adapted to them;
- a strategy of mutual adjustment. In this strategy both sides are developed at the same time. Required qualifications and qualifications at hand are adjusted to each other in on and the same process of working and learning.

We expect these strategies to be applied in differing degrees in different types of organisations. Hypothetically, organisations use different strategies for reducing qualification problems and this expresses itself in different learning-strategies. Different organisations create different learning opportunities for their employees and organize them in a different way. The supply of learning programmes and the organisation of the learning function vary between organisations. This variation manifests itself very clearly in three types of organisation Mintzberg (1979) calls the machine bureaucracy, the professional bureaucracy and the adhocracy.
- In the machine bureaucracy a task-oriënted supply of learning programmes is to be expected. Learning programmes are developed on the base of an analysis of tasks, jobs and functions within the organisation. Trough these programmes the qualifications of the members of the organisation are adapted tot the (new) task- and job-requisites.
- In the professional bureaucracy a profession-oriënted supply of learning programmes is to be expected. The qualifications of the organisation members are adapted to the developments in the field of the profession. Besides formal professional education and training outside the organisation much emphasis is laid on learning on the job.
- In the adhocracy learning and working are strongly integrated. Employees develop their qualifications primarily in and through their work.

This paper focusses on learning in professional organisations. A (typical) professional bureaucracy is primarily an operational organisation. There is almost no technical staff and there are only a few managers. The core of the organisation are the professionals, conducting the primary processes. A strong side of this type of organisation is that the professionals can use their professional qualifications directly in practice. Tasks and functions within the organisation strongly correspond with what they were trained in during their period of education. The work of the professionals within the organisation is structured conform the structures of their profession.
Professional bureaucracies react very slow on changes in their environment, for instance on changes in target-groups or changes in questions and problems of cliënts. This is related to the low level of structural provisions for developing new products or services. These provisions are not situated inside but outside the organisation. An essential condition for professional bureaucracies is that outside the organisation there are special agencies for development and education in the professional field. The profession has to function well. In a professional bureaucracy we expect a profession-oriented learning-strategy to be dominant.

1019

This strategy typically has the following characteristics:

- the supply of learning programmes is profession-oriënted; it is based on (scientific) developments in the profession and intends to adapt qualifications of professionals to new developments in the professional domain;
- the most important supply comes from outside the organisation, in the form of courses and seminars, offered by the professional association, in which new developments in the field are presented;
- an important part of learning also takes place on the job; by means of self-study, internal consultation on work and study within the discipline and evaluation of their own performance professionals keep up their routines and develop new ones;
- the professional himself is responsible for his learning activities, i.e. for his participation in the courses and seminars of the professional association;
- the management of the institution primarily facilitates qualification-development of the professionals by creating adequate conditions in the sfere of time, money, practices of internal consultation and work evaluation.
- the training staff, as far as present, primarily has a (limited) intermediary function.

Research questions and research design

Our basic question is the following: to what extent do professional organisations indeed apply a profession-oriënted learning-strategy? In order to answer this question a comparative casestudy design is used. We conducted casestudies in four professional social service institutions in the field of care and rehabilitation of offenders and (ex-)prisoners. In gathering the data a multi-method approach was followed. In each case information on learning programmes and on the organisation of the learning function was gathered by an analysis of documents, a series of interviews with management, staff and leaders of operational professional groups and group discussions within the context of procedures of datafeedback. The data were analysed with within-site and cross-site matrix-techniques.

Results and discussion

In the investigated institutions our expectations considering the profession-oriënted learning-strategy partially came true. In all four institutions the profession-oriënted strategy is indeed of major importance, but it is applied rather one-sided. Especially the opportunities for learning within the organisation are developed and situations of learning on the job are frequently found. This has much to do with the organisation of work in the four institutions. Many learning-opportunities are provided within the professional teams, the basic units of the organisations. Grossly, teams are organised according to the main division in the field. In all institutions there are teams for extramural and intramural care and rehabilitation; in two institutions there are also separate teams for new specialisations, i.e. social-psychiatric work and alternative sanctions. Within these teams, each professional is an allround worker. He handles all types of (new) cliënts and all types of problems and he applies all types of methods and programmes in use. His work offers him a lot of individual learning-opportunities, both through the contacts with his cliënts as also through the contacts with his collegues and with collegues of other institutions. Besides that, many groupwise learning-opportunities are created within the teams. Most teams organise workload- or caseload-meetings, during

1020;

which specific cliënts and possible ways to deal with them are discussed. Most teams also organise some kind of intervision, meetings during which team-members reflect on their behaviour toward specific cliënts or specific problems. In most teams study-sessions are held around actual themes, like incest-crimes or agressive cliënts. If new types of problems gain high priority special task- or projectgroups are formed, in which the new problems and new methods to handle them are explored by members of different teams. By all these means the professionals try to keep up and renew their routines and qualifications. The teams are highly autonomous in organising their learning-activities.
While on the job learning within the teams is well developed, off the job learning is not. Courses outside the organisation in which new developments in the profession are presented are rarely followed. This, we think, is a consequence of the low level of professionalisation of the occupational sector. There is no clearly defined occupational domain, no specific infrastructure for initial education and further professional training, no well-organised professional association and not much research and organised reflection in the professional field. In all four institutions on the job learning is the dominant strategy of professional qualifation-development. By lack of an educational infrastructure there is no off the job supply.
In spite of this similarity, we also found clear differences in qualification-policy between the four institutions. Organisations in the same service-field obviously react differently on developments in their environment. Explanations for these differences are to be sought in the semantics and social relations of management and professionals as the central actors in the organisations.

- In one institution qualification-policy is highly formalised, including instruments of needs-assesments, planning, budget-allocation and evaluation; there is a special training-staff. Here management gives high priority to qualification-development, a.o. as an organisational instrument for coping with new developments. In two other institutions formal qualification-policy is primarily a set of rules regulating individual professional requests. In the fourth institution qualification-policy is not yet formalised. Here management gave priority to a process of reorganisation in order to prepare the organisation for its expected new tasks.
- In two institutions the internal learning efforts mainly concern substantial professional themes (new routines, new problems, new methods). In these institutions management is strongly profession-oriënted and from this oriëntation it firstly tries to tackle changes with a strengthening of professional qualifications. In the other two institutions larning efforts mainly concern the outcomes of the professional services and the organisational aspects of the professional work-routines. Here an organisation-oriënted management-style is dominant. Management organises evaluation-sessions of services and routines or policy-groups for discussing elements of the institutions policy and thus tries to develop besides professional also some organisational qualifications among the otherwise autonomously operating professionals.

STACKING AND TINNING: HOW STEELWORKERS LEARN THEIR JOBS

J.A. DeJong, I. Pool, & G. Van Elst, Department of Education, Utrecht University, Utrecht, The Netherlands

Only a few evaluation studies on systematic on-the-job training have been reported up to now (Belbin et al., 1957; Cullen et al., 1976; Jacobs, 1991; Jacobs & McGiffin, 1987; Jansen, 1990; Kolkhuis-Tanke, 1990). This paper is a report of an evaluation study of two job training programs at Hoogovens IJmuiden, a large steel works (blast furnaces) in the Netherlands. Job training at Hoogovens is structured according to a format, which is developed in the Hoogovens plants. This format, and its development, are sketched with reference to Reiss (1991) . An interesting feature is that the format of job training used at Hoogovens was not developed by the training department, but by local management on the shop floor. Part of the training takes place in the form of structured training on-the-job and part in the form of off-the-job courses. Structured training on-the-job is defined as: *...a task-oriented training (or part of such a training) that takes place on the shop-floor. The learner carries out assignments according to a written plan, being coached by an experienced colleague or by a supervisor. At the end the mastery of the tasks is formally assessed*".

Our evaluation is concerned primarily with the on-the-job part.

The study presented in this paper is part of a larger research program on the subject of on-site training. We sketch here the theoretical embedding of the study, and an outline of previous and forthcoming studies within the program.

The research reported took place in a plant where packing steel is produced which is used for tin cans, crown caps, and battery cases. Thin steel plates are electrolytically covered with a layer of tin, and then cut and stacked. The training programs studied were a job training for blue collar personnel, who stack the plates (referred to as 'stackers'), and a job training for supervisory personnel, who control the tinning ('tinners'). Documents were studied, and trainees, instructors, supervisors and training co-ordinators were interviewed (N=38) in order to find out how the training as intended works out on the shop floor.

The evaluation is based on a comparison of the program-as-realized with the program-as-intended (compare Stake, 1969). The 'history' of the programs was divided in four phases: initiation, development, operation, effect. Data were gathered on intentions and actualities with regard to all four phases.

1022

In this paper results on the phases '*operation*' and '*effect*' of the two programs are presented and discussed. The following themes will be dealt with:

- the relationship between training and career-decisions;
- the timing of off-the-job courses;
- the sequence of on-the-job 'modules';
- instructors;
- the use of handbooks;
- the use of checklists;
- coaching styles;
- learning styles;
- assignments;
- competency tests;
- quality control;
- monitoring the system.

They themes are illustrated by means of excerpts from the interviews.
The following are some conclusions of the investigation:

- An important advantage of on-the-job training is the possibility of 'training just in time'. This timing factor appears to be a very important mediator of the effect of a job training program. Our study shows good and bad examples of timing. Much depends on the way supervisors manage the tension between production and training. Timing appears to be even more difficult for off-the-job courses. In the majority of the cases these courses come either too soon or too late. Individual instruction (for example, computer aided instruction) seems to be part of the answer to this problem.
- Handbooks with process descriptions and work prescriptions prove to be a valuable aid for on-site instruction. The handbooks are regularly used. Workplace instructors use them to support their explanations and learners use them to explore their job. One problem with these handbooks is their currency. It requires sophisticated logistics to keep them up to date.
- An interesting observation is that supervisors attach great value to certain learning style characteristics (especially initiative) of their new personnel.
- Linking training results with career development turns out to be an advisable policy. It increases motivation for training and learning, as well as the objectivity of career decisions. Unwanted side effects can occur, however. This is the case if training obligations are treated in an overly formal manner.
- For several reasons evaluation of the training by measuring quality of products turns out to be difficult to effect. Subjective confidence of local managers seems to be an evaluation criterium that suffices for most persons concerned.

The present study, supplemented with other studies, makes it possible to formulate some elements of a theoretical model for the description of on-site training.

References

Sloman, M. (1989). On-the-job training; a costly poor relation. Personnel Management, February 1989, 38-41.

Belbin, E., Belbin, R. M., & Hill, F. (1957). A comparison between the results of three different methods of operator training. Ergonomics, 1 (1), 39-50.

Cullen, G., Sawzin, S., Sisson, G. R,. & Swanson, R. A. (1976), Training, what's it worth?; an experimental case study at Johns-Manville Corp. Training and Development Journal, 30 (8) 12-20.

DeJong, J. A., VanElst, G., & Pool, I. (1991). Werkplekopleiden bij Hoogovens IJmuiden. Utrecht: Rijksuninversiteit Utrecht; Vakgroep Onderwijskunde.

Jacobs, R. L. (1991). Forecasting the financial benefits of unstructured and structured OJT. Paper presented at the International Conference on Research on Corporeate Training and Development at Enschede (Nl.), Sept. 27, 1991.

Jacobs, R. L., & McGiffin, T. D. (1987). A human performance system using a structured on-the-job training approach. Performance and Instruction, july 1987, 8-11.

Jansen, M. (1990). Evaluatie opleidingen monteur en hoofdmonteur seinwezen; werkplekopleidingen getoetst aan de praktijk. Utrecht: Rijksuniversiteit Utrecht; Vakgroep Onderwijskunde (Master thesis).

Kolkhuis-Tanke, I. R. (1990). Evaluatie proefopzet: flexibele opleidingsorganisatie Lijnwerkplaats Amsterdam (unpublished report). Utrecht: Nederlandse Spoorwegen.

1024

COMPANY TRAINING: VOLUME, UNDERINVESTMENT, AND RETURN

J. de Koning, & A. Gelderblom, Department of Labour-Market and Education, Netherlands Economic Institute, The Netherlands

Introduction

Company training is assumed to be economically valuable. Because of the direct involvement of companies, the assumption is that company training, more so than other types of training, satisfies economic needs and thus brings in a high economic return. On the other hand, in the literature the training market has been pointed out to suffer from imperfections which allegedly cause underinvestment in training, especially in general training. A company would leave it to the employee to pay for his general training, lest a rival should profit from the company's investment by "buying" its trained employee. The employee would also be reluctant to invest in - costly - training because of the uncertainty of the private return. So far, however, little has come to light about the return on company training and the extent of underinvestment. The present paper, after considering the economic significance attached to company training and sketching a theoretical framework, presents the results of two investigations we have recently conducted into company training.

Nature of the investigations

The first investigation was based on in-depth interviews with over 50 companies, chosen from six sectors, in which the following aspects of company training were discussed: volume, nature, costs, distribution among groups of employed, the role of the environment (market structure, collective training agreements), the extent of underinvestment within companies, and the effects on company performance. Underinvestment was measured in two ways:

- by questions about possible bottlenecks or obstacles in enlarging the training effort;
- by asking a direct indicator for the discrepancy between wanted and realised training effort.

Both indicators turned out to be strongly correlated, which indicates that both can be seen as a reliable way of measuring underinvestment.

The effect of training was measured by a two-equation model in which the training effort and the productivity of the firm are explained by each other and some control variables.
The second investigation was based on data from a large-scale survey among companies which was commissioned by the (OSA, Agency for Strategic Labour-market Research) in 1988. This survey covered a wide range of aspects, among which technology, internal organisation, production volume, composition of the staff by such characteristics as age, educational level and sex, wage structure, work conditions, participation in company training and financial data. Using regression analysis we focused on the question of what factors explain the productivity. What role have, among other factors, the participation in internal and external training in answering this question?

Results

In confirmation of the results of other research, the participation in training courses appears to be higher if the company is larger. Small companies appear to make frequent use of external courses, while large companies tend to make use of courses which are exclusively given to their own employees. On average, about half of all company training courses are internal courses. About two thirds of the training courses can be considered as general training. The majority (over three fifths) are organised by private institutions. Strikingly, the discrepancy between wanted and actual training efforts is not wider among small than among large companies, despite the lower level of company training in the small companies. Nor do small companies suffer more bottlenecks in their training investment. Small companies probably compensate their lower level of formal training by more on-the-job training, or have no clear view of the advantages of formal training. Another remarkable result is that companies receiving the most subsidies for training show the widest discrepancies in education. Apparently, collective agreements and subsidies, though developed to overcome training discrepancies, have only limited effects on that score.
In both investigations the connection between level of education and productivity of the employed has been tested econometrically, as mentioned before. The possibility of mutual influences between the level of education and productivity has been taken into account. Because the investigation seems to suggest that to raise productivity is an important motive for company training, a causal relation from training to productivity has been assumed. On the other hand, productivity can influence training efforts: well-performing companies have indeed more financial scope to invest in training courses. The analysis has also taken into account that other factors may contribute to the explanation of both variables. The composition of the staff and the position of the trainers within the organisation are two examples of explanatory factors for training efforts, while for example the technology used by the company is an important explanatory factor for productivity.
The results point towards a positive if small effect of company training on productivity. External courses have more effect than internal ones. That the effect of company training is small is not surprising, for two reasons:

- because informal training has not been taken into account, the effect represents in fact the surplus value of formal training in comparison with informal training;

- not all companies taking part in the interviews have adopted a structured training strategy. Mostly they have a poor idea of the returns and know even less of the costs of company training courses. Therefore we must assume that they are not making optimum use of the training instrument.

Points for further consideration

Certain aspects uncovered by the investigation merit closer attention. One question is how far small companies compensate their smaller investments by more on-the-job training. Is such training a suitable substitute for formal training? If so, our assessment underestimates the effects of company training. Another point to be raised is the evaluation of training courses by the companies themselves. Indeed, many companies hardly plan or evaluate their training efforts. Although they claim to be interested, they are scarcely able to visualise the effect of training. How far can existing knowledge of evaluation methods serve to make practical recommendations to companies about how to evaluate their company training?

References

Becker, G.S. (1964). Human Capital. New York.

Bishop, J.H. (1988). Do Employers share the Costs and Benefits of General Training?. Cornell University.

CBS (1988). Bedrijfsopleidingen in Nederland in 1986 (Vocational training financed by the private sector in the Netherlands in 1986). Den Haag.

Feuer, M., Glick , H., & Desai , A. (1987). Is Firm-Sponsored Education Viable? Journal of Economic Behaviour and Organization, 8, pp. 121-136.

Gelderblom, A., & Koning, J. de (1992). Meer-jarig, minder-waardig? Een onderzoek naar de invloed van leeftijd op produktiviteit en beloning (A study of the influence of age on productivity and wage). Den Haag.

Kodde, D.A. (1987). Uncertainty and the Demand for Education. The Review of Economic Statistics, pp. 460-467.

Koning, J. de (1991). Evaluating training at the company level, paper for the International Conference on the Economics of Training, Cardiff, 23-24.

Koning, J. de, Gelderblom, A., Hammink, A., & Olieman, R. (1991). Bedrijfsopleidingen: omvang, aard, verdeling en effecten (Company training: volume, character, distribution and effect). Utrecht.

Ritzen, J.M.M. (1989). Market Failure for General Training and Remedies, paper for the Robert M. La Follette Symposium on Market Failure for Training, 11-12.

LEARNING AND TRAINING STYLES IN CORPORATE TRAINING

J. van Zuilichem, & J.D.H.M. Vermunt, Department of Psychology, Tilburg University, Tilburg, The Netherlands

Philip Morris Holland, a modern high-tech industrial plant, expressed us the wish to develop more into the direction of a 'learning organization'. An important aspect of such an organization is the independent learning capability of the employees. We were asked to develop ideas about a training concept that more closely links to the characteristics of a learning organization than their current training practices. At Tilburg University we recently developed an instructional design theory called 'process-oriented instruction'. It is defined as instruction directed at initiating, supporting and influencing the learning and thinking activities students use to learn. Learning content and learning activities are taught in coherence. Its ultimate aim is to promote students' skill in selfregulated learning and problem solving. The theory fits within the constructivist movement that recently arose in instructional psychology and design. A very important first step in designing instruction is, according to this process-oriented theory, a thorough diagnosis of the current learning and training practices. Therefore we decided to perform first a diagnostic study on the learning and training styles prevalent in the organization. The study was aimed at answering the following questions:

- What learning styles are characteristic for course members of Philip Morris?
- How are these learning styles associated with several personal and contextual factors and with evaluations of courses?
- What training styles are characteristic for trainers of Philip Morris' courses?
- How are these training styles connected to several personal and contextual factors?
- What are the relations between training styles and learning styles?

Two diagnostic instruments were constructed, based on the Inventory of Learning Styles (ILS) developed earlier at Tilburg University. The *Questionaire Learning Styles (QLS)* was intended to measure the extent to which course members used a variety of cognitive and metacognitive activities when learning. Besides their learning conceptions and orientations were measured with this instrument, as well as background variables and course evaluations. The *Questionaire Training Styles (QTS)* was intended to measure the extent to which the trainers of this course members capitalized on, substituted, or activated all cognitive and metacognitive activities that were part of the QLS. Also questions about background data were asked. The QLS was administered to 259 course members

1028
201

of seven different course areas, the QTS to 30 trainers of the same course areas. Data were analyzed via reliability, correlational, variance and regression analyses. Concerning the first question the results indicated that four learning styles could be distinguished: a meaning directed, a reproduction directed, an application directed and an undirected learning style. With regard to the second question it turned out that the learning styles of course members were most related to prior educational level, functional level, course area, number of years ago that the prior education was completed, and number of internal courses followed. There were only a few relations between learning styles and age, number of years working at Philip Morris, and number of functional changes. The results also showed that learning styles were associated with the way course members evaluated the difficulty, pace, adaptation to prior knowledge and learning effects of courses. The main result with regard to the third question was that training styles are more accurately described in terms of meaning directed, reproduction directed, application directed and undirected styles than in terms of capitalizing, substituting or activating styles. These training styles were found to be most associated with years of training experience and number of training hours. The more experience trainers had, the more meaning directed their training style tended to be. The more hours a week trainers gave courses, the more reproductive their training style was. The analyses carried out to answer the fifth question revealed some very peculiar relations. In general, the training activities of trainers were less associated with the learning activities of their course members than we had expected. Interestingly however substituting training activities were more related to the learning activities course members employed than activating training activities.

Theoretical implications of this study have to do with the gained understanding of the nature, function and correlates of learning and training styles in corporate training. The description and operationalization of learning and training styles based on constructivist learning theory proved to be a powerful way of diagnozing current practices. The study revealed more insight into the presence and absence of interplay between learning and training. Practical consequences have to do with the revelation of strong and weak points in current training practices compared to the characteristics of a learning organization.

THE 'NEW FLEXIBILITY' IN VOCATIONAL EDUCATION

David Raffe, Centre for Educational Sociology, University of Edinburgh, Great Britain

General issues: the pursuit of flexibility in education systems

Pressures on vocational education

Vocational education in all countries is subject to similar economic, social, governmental and individual pressures. Economic pressures: education is required to increase skill levels to promote economic competitiveness; it must supply the new skills required by a changing economy; and it must prepare students for a labour market characterised by occupational uncertainty and mobility, and by high unemployment. Social pressures: vocational education must attract students when aspirations are rising; it must cater for students with increasingly diverse backgrounds, interests and needs, including the unemployed, ethnic minorities and other disadvantaged groups; and it is expected to promote equality of opportunity. Governmental pressures: education is accountable to a wider range of interests, including students, parents, social partners and the 'community'; it must handle increasingly complex relationships between local, regional, national and supra-national levels of government. Individual pressures: education is increasingly expected to empower the individual and to give him or her 'ownership' both of the learning process and of the skills it provides. More demands are being placed on vocational education; future demands are increasingly unpredictable; and these demands must be satisfied in conditions that are also uncertain.

The response: flexibility

Most systems are responding to these pressures by attempting to increase their flexibility. There are several dimensions of flexibility in education, including: individual flexibility as a competence acquired through vocational education; curricular flexibility, both in the sense of easy updating and in the sense of being able to tailor programmes to respond to local circumstances and to meet the varied needs of individual students; flexibility of delivery, so that the same outcomes can be achieved by learning in different institutional contexts, by different methods or over different time periods; and flexibility of pathways, which involves more open access to education, weaker divisions between categories of students, and more diversity in the occupational or educational destinations to which each pathway may lead. Measures to increase flexibility include: the decentralisation of decision-making; closer links with the 'world of work'; curricular changes to increase

1030

breadth, defer specialisation and extend student choice; credit systems and arrangements for transfer between educational tracks and specialisms; 'open' learning and wider access arrangements; encouragement for more varied teaching and learning styles, including work-based learning; and greater institutional diversity.

The problem of coordination

Flexibility is primarily a characteristic of systems, and an effective flexible system requires coordination. It therefore faces a problem: how to provide coordination and maintain coherence when sources of control and initiative are dispersed. Countries adopt a variety of solutions to this problem. In this paper I pay particular attention to the use of vocational qualifications (in the British sense, of certificates) as a means for coordinating systems.

An economic analogy

There is a rough analogy with countries trying to move from a command economy towards a more flexible market economy. One could identify at least four types of 'economies' within education:

- Command economy;
  control is centralised and inflexible.
- Barter economy;
  flexibility is achieved through local linkages (for example, between schools, colleges and employers) but there is no central coordination.
- Monetary economy;
  flexibility is achieved through decentralised decision-making, coordinated through a 'currency' of qualifications. There are two variants:
  soft currency: flexibility is achieved within a restricted range (the 'currency' of qualifications is only convertible within a given locality or sector) and
  hard currency: flexibility is achieved across the whole system.

There are three conditions for a 'hard currency' economy: the currency of qualifications must be fully convertible (that is, recognised throughout the system regardless of where the currency has been issued); there must be effective central control of the money supply (that is, of the introduction of new qualifications); and most importantly, the qualifications must have exchange value - they must buy things that people want to buy.

Case study: the pursuit of flexibility in Scotland

Context

The economic, social and individual pressures for flexibility in Scotland are similar to those in other countries. The governmental pressures may require more explanation. Scotland has its own education system, with separate policy-making for education, but policy for industrial training is made for the whole of Great Britain. Scottish education must therefore respond to changing Great Britain training policies while retaining coherence as a (Scottish) system. (The need to respond to EC training policies creates similar pressures for flexibility in member states.) The Scottish education and training system is a 'mixed model', in which 16-19 year olds can choose between different sectors: school; full-time college

courses; youth training (YT) schemes (which often include part-time study at college); and other work-based or part-time study usually sponsored by employers. There is an additional pressure, therefore, for flexible coordination of these sectors.

The Scottish model of flexibility

The Scottish 'strategy' for flexibility is based, at least implicitly, upon these four principles:

- The diversity of the 'mixed model'. A diverse system (it is claimed) can more effectively cater for diverse student needs; moreover, the sectors closest to industry can respond flexibly to changing economic demands;
- Modularisation. All vocational education was modularised in 1984. The modules are notionally 40 hours in length, competence-based, criterion-referenced and certificated by a single qualification (the National Certificate, or NC). Modules are intended to make it easier to update the curriculum, to extend choice for students and other customers such as employers, to create more flexible pathways and to permit more choice of the method, location and pace of study. At the same time as NC modules were introduced, there were reforms of curriculum, pedagogy and assessment.
- A single 'currency' of qualifications. The modular system is national, and covers all sectors of education (school, college, YT, etc). NC modules from the same catalogue may be taken in different sectors; students moving between sectors may follow progression routes within the modular framework. NC modules are defined in terms of 'learning outcomes' and their value should not depend on the pathways or methods by which these outcomes are achieved;
- Flexibility between general and vocational education. NC modules may be used to deliver a 'general' curriculum, especially to lower-attaining school students. In addition, Highers - the main 'academic' school courses for 16-18 year olds - are short (one-year) courses leading to (ungrouped) single-subject qualifications; it is therefore easy for school students to combine Highers with NC modules.

The Scottish system in action

Six in ten of a recent Scottish year group had started at least one NC module by the age of 19; participation is rising. There is no 'typical' NC student; different types of young people study modules for different purposes. Modules are studied at school, in full-time college, on YT and in employment; there are differences between these sectors in the backgrounds and characteristics of NC students, in the numbers of modules studied, in the amount of choice and (to some extent) in the subjects studied. The take-up of modules by school students has been unexpectedly high; by 1989 more than half of 16/17 year olds at school were taking modules, usually together with academic courses. Many young people take NC modules in more than one sector - for example first at school, and then at college or on YT - and these students tend to continue in the same subject areas. There is a fairly close match between NC subjects and the occupations of young people. Some NC students have progressed to higher education. Students have expressed favourable attitudes to the curriculum, pedagogy and assessment of NC modules, although modules are perceived to have low status both in the education system and outside it.

1032

The Scottish modular system has therefore responded reasonably well to the diverse demands of the 'mixed model', and it has increased the coherence of the system. 'Packages' of modules have been provided for a variety of different purposes, and these packages can be changed flexibly in response to new demands - including new training policies introduced from outside Scotland.

Limitations and problems

However the Scottish experience raises a number of issues, some of which may be common to other countries pursuing flexibility.

- Academic drift. The system is tending to become top heavy, with too few students pursuing intermediate level qualifications. This may reflect the system's 'flexibility', with weak boundaries between academic and vocational study; or it may reflect insufficient flexibility, because the pathways leading back from vocational to academic study are too weak.
- Issues of choice and flexibility in a modular curriculum. There have been problems in providing appropriate options within the modular curriculum, in guiding student choices, in defining clear pathways and in ensuring that the curriculum is coherent. There has been a partial retreat from the principle of 'modularity'.
- Issues of co-ordination in the 'mixed model'. Modules awarded in one sector (eg school) are not always recognised in other sectors (eg college or YT); there are many sources of initiative in the system, so the total number of modules tends to grow too large (there are now 3000); qualifications perform different functions in different sectors of the mixed model, and no single system of qualifications can perform all these functions equally well.
- Limited effects on participation. Participation in the NC has increased among young people within different sectors (school, YT, etc); but the flexible modular framework has not significantly changed total participation in each sector, or the types of students who enter each sector. Participation is still too low, and traditional inequalities (for example of gender) persist.

The failure to establish a 'hard currency' of qualifications. The paper argues that these limitations and problems arise (partly) because Scotland has not fully achieved a 'hard currency' of vocational qualifications. The three conditions listed earlier have not been met: the qualifications have not been fully convertible - their value has varied according to where they were obtained and where they are being 'spent'; the 'money supply' - the development of new modules - has run out of control; and most importantly, the modules have had low status and limited exchange value in either educational or labour markets.

Implications for research

I conclude by discussing the implications of the 'new flexibility' for educational research.

NEW METHODOLOGIES FOR IN-COMPANY VOCATIONAL TRAINING AND EDUCATION: EMPIRICAL CASE STUDIES

Brigitte Schmidt-Hackenberg, Bundesinstitut für Berufsbildung, Berlin, Germany

The German system of vocational education and training through apprenticeship is conducted and financed by the companies themselves but subject to public responsability and control. The Bundesinstitut für Berufsbildung, Federal Institute for Vocational Education is promoting in-company vocational training and education through research, counselling and in behalf of the Bundesminister für Bildung und Wissenschaft (Federal Ministry of Education and Science), through funds. The advancement of training methodologies is one of several areas of public funding for pilot projects of in-company training. The funding covers the larger part of the additional costs incurred by developing and testing new training concepts.

The pilot projects on new training methodologies were conducted in the fields of metalworking, electrical engineering, commercial training and the painting and lacquering craft.

The pilot projects differed considerably in terms of genesis, objectives, motives and concepts; the notional diversity being confusing, it seemed that formalized rather than substantive categories were better suited for describing and evaluating the pilot projects:

- Need for innovation: What circumstances called for change in the training philosophy?
- Innovation approach: Which concept was selected? For what reason?
- Introduction strategy: How were the persons affected prepared for the changes?
- Implementation: How was the concept implemented, how modified during implementation?
- Evaluation: What was the task of the flanking scientific monitoring?

The analysis is based on preliminary application dossiers, interim reports and final reports, on personal observations and oral accounts.

By emphasizing self-reliant planning, execution, monitoring and control of ones work the pilot projects have been of some influence in changing educational standards of in-company vocational training in Germany.

Only some of the projects can be mentioned here. A joint project "Motorsegler", a sailing ship with auxiliary motor, was used in the training of miners; the "Lehrlingsbaustelle", a construction site for apprentices only to train them as painters and lacquerers. The famous "Dampfmaschine", a model steam engine, and all that it incorporated in the pilot project of the Daimler-Benz AG, Gaggenau,

1034

made the company a "Mecca" for vocational training. Equally interesting are the follow-up projects for the specialized (2nd and 3rd year) training. The method of "Leittexte", guidance scripts developed at the Stahlwerke Peine-Salzgitter AG, Salzgitter consists essentially of written guideline questions worked out by the trainers themselves to encourage the trainies to more independence and selfreliance. The training concept of the J.M. Voith GmbH at Heidenheim, successful since long and well known among experts, is based on the compensatory effects of creative work to crafts and technical training.
The pilot projects differ widely as to the role of scientific theory and research. Theories of learning gave the impulse to the concept of guidance scripts. But other successful projects have been initiated without previous scientific analysis. Scientific monitoring in one case influenced a project substantially, in others it was merely describing proceedings and results.
As a consequence of the pilot projects psychological theories of action which the method of guidance scripts is based on, have gained much attention in the pedagogy directed to technical and commercial education. In these theories acting and thinking are regarded as fundamentally intertwined, thus giving to the practical vocational learning a new theoretical legitimization and new pedagogical selfconfidence.

References

Schmidt-Hackenberg, B., Höpke, I., Lemke, I.G., Pampus, K., & Weissker, D. (1990). Neue Ausbildungsmethoden in der betrieblichen Berufsausbildung. Ergebnisse aus Modellversuchen. 3. Printing with supplements in English, French and Spanish. Berlin: Bundesinstitut für Berufsbildung. (Berichte zur beruflichen Bildung, Heft 107).

APPRENTICESHIP MODELS IN THE METAL INDUSTRY: FIGURES AND VISIONS

A.F.M. Nieuwenhuis, & J.A. Wiggers, RISBO, Erasmus University Rotterdam, Rotterdam, The Netherlands

Context

In this paper we present a research project on the training system for craftmanship in the metal industry. The Dutch metal industry consists of two branches: the metal-electro industry (mainly large companies) and the metal-working industry (mainly small and mediumsized companies). We looked at the training system for jobs and functions like welding, fitting and construction engineering. For these jobs the apprenticeship system is the main road to becoming skilled.

The apprenticeship system in the Dutch metal industry cannot be characterized by a single model for on-the-job training. After the economic recession around 1980, several models are developed to ensure a regular flow of new workers for the two branches of industry. Nowadays three main models are distinguished in the collective labour agreements: the A-, B- and C-model. The A-model consists of practical learning in a pure off-the-job situation, the C-model consists of purely on-the-job learning, whereas the B-model forms a mixture of those two. Within these main models large differences can occur: the off-the-job training can be located in training centres outside the companies, company-schools or practice corners, whereas on- and off-the-job learning can be combined with practical periods in other companies, either on- or off-the-job.

From an instruction-technological point of view it is interesting that the off-the-job models are mostly used by large companies in the Metal Industry, whereas the on-the-job models are preferred by small and medium enterprises. The possibilities to learn different skills on-the-job are smaller in large companies because of a higher degree of automation. Activities can be varied more in mediumsized companies, especially if they work on a craftsmansway. Small companies use the C-model, but have difficulties in offering varied learning opportunities. Large companies employ more trainees, so they can organize efficiently company-schools. Apperently the preferences for the different models are also induced by other than instruction-technological arguments: efficiency, lower wages and a shrinking labour market are arguments which are heard often.The use of different models causes many unjustices and frictions between the two branches of industry and also between employers and trade unions. The smaller enterprises use the training facilities of the larger ones or hire trained workers without any own training activity. The work and wage contracts of apprentices differ largely between the different models,

1036

which is unacceptable for the unions. For these reasons the training funds of both branches commissioned a research project to clarify the pros and cons of the used models and variations.

Design

We designed the research project as an evaluability assessment (Rutman, 1980): our main focus was to discover the mean-end argumentations which are used by educational policy makers in the branches; the second target was to make an inventory of possible benefits (and disadvantages) of the models both in educational as in labour market terms. Therefor we made a theoretical analysis based on recent instruction-technological literature (Raizen, 1989, 1990; Achtenhagen, 1989, 1991; Nieuwenhuis, 1991; Simons, 1991), we collected available figures from the national training foundation for metal jobs (SOM) and we interviewed experts on branche level. Within 31 companies and training centres we interviewed 25 employers or managers, about 35 trainers and about 40 apprentices. We selected the companies in such a way that the different models and variations are represented equally and that some variation in branches, regios and metal jobs is assured. In the triple interviews we used the model of cognitive apprenticeships as structuring theory (Collins a.o., 1989).

Some theoretical notes

The different apprenticeship models used for the training of metal craftsmen can been placed on a scale from authentic to simulated learning situations, within one or more companies. Frietman (1990) concludes in his thesis that the different models are equally adequate training models. He prefers the C-model, because this is the traditional apprenticeship model. Simons (1991) disagrees with Frietman: based on a secondary analysis of Frietmans data he argues that mixed models are more adequate than the traditional model. Transfer of skills to new work situations plays an important role in Simons statement. Also Nieuwenhuis concludes in his thesis (1991) that on-the-job situations are not automatically the best training situations. He states that the 'best' learning situation does not exist: a good training model consists of a sequence of different training situations.
For the training of metal workers one can conclude that a good training sequence goes from model A via model B to model C. In the first part of the training technical skills can be stressed on whereas in the latter part of the training workexperience and social and complex skills can get more attention. Decisions on the shifts from one model to another depend on the learning skills of the trainees.

Visions of policy makers

Based on the interviews on branche level we can conclude that the policy makers see benefits in all training models. They agree on the idea that different models should be combined to create adequate training courses. Their visions are not only based on instruction-technological arguments, but also (and mainly, ?) on institutional goals: off-the-job training systems are less vulnerable for economic developments, can deliver better training for non-traditional groups (for example girls and ethnic minorities) and can also been used for retraining and reskilling of employed workers. The skills and motivation of the trainers have a large impact on the quality of the training, just like the training facilities in the different situations.

1037

For the future local and regional training facilities and policies are necessary to garantee a well skilled workforce for the metal industry.

First results from the local interviews

From the interviews on location we got the impression that everybody is enthousiastic for his own training model. Each model can be a good model if one is willing to invest time and money. The more pure a model (off- or on-the-job) the less adequate: mixed models can garantee good training. Learning on-the-job forms an essential part of the training but trainees should also get the possibility to make mistakes and to practice skills. Company-schools and training centres are more adequate for the training and the recrutement of non-traditional trainees. Training the trainers should be important in future policy. More results will be presented at the conference.

Preliminary conclusions

Instructional theory, policy visions and local strategies and results seem to merge together: a mixed and flexible training model offers the best garantees for a well skilled and learning workforce for the metal industry. Employers and trade unions should create the facilities to build such a training system.

References

Achtenhagen, F. (1989). Development of problem solving skills in natural settings. Key-note adress for EARLI-conference Madrid, 1989. (Will be publiced in the conference proceedings).

Achtenhagen, F. (1991). Learning, thinking, acting in complex economic situations: models of preparation for vocational and private life. Paper for the 4th EARLI-conference in Turku, Finland. Göttingen BRD: Georg August Universität.

Collins, A., Brown, J.S., & Newman, S.E. (1989). Cognitive apprenticeship: teaching the crafts of reading, writing and mathematics. In: L.B. Resnick. Knowing, learning and instruction. Hillsdale, New Yersey: Lawrence Erlbaum Ass.

Frietman, J. (1990). De kwaliteit van de praktijkcomponent in het leerlingwezen (diss). Nijmegen: KUN/ITS

Nieuwenhuis, A.F.M. (1991). Practical learning situations as a preparation for lifelong job oriented learning. Invited paper for the international workshop on developing education for lifelong learning in Tampere, Finland. Rotterdam: RISBO/EUR.

Raizen, S.A. (1989). Reforming education for work: a cognitive science perspective. Berkeley, CA.: national centre for research on vocational education.

Raizen, S.A. (1990). Learning and work: the research base. Discussion paper for the seminar on linkages in vocational-technical education and training. Phoenix, AR: US state department of education/OECD.

Rutman, L. (1980). Planning useful evaluations. Evaluability assessment. Sage Publications.

Simons, P.R.J. (1991). The promotion of transfer of learning to work-settings. Paper for the 4th EARLI-conference in Turku, Finland. Nijmegen: KUN.

1038

EVALUATION OF APPRENTICE-SYSTEM COURSES IN INFORMATION SCIENCE

M.S. Boekenoogen, & J.W. Eberg, ISOR, Department of Educational Research, University of Utrecht, Utrecht, The Netherlands

Context and research questions

From 1986 until 1991 the Educational Research Department of the University of Utrecht (Holland) carried out two research projects for the evaluation of information-science courses in the Dutch apprentice system. As commissioned by the National Centre for the Innovation of Vocational Education and Training (first known in its abbreviated Dutch form as PCBB, now as CIBB) the experimental ISI courses were evaluated in the period between 1986 and 1989 (see Stokking & Eberg 1989). Subsequently, research into four new regular information-science courses, which had been commissioned by the Dutch Vocational Training Centre for Economics and Administration (in Dutch: ECABO), was conducted from 1989 until 1991 (see Boekenoogen, Eberg & Stokking 1991). The aim of both research projects was to gather data during the implementation of the courses and to structure these data in such a way that they could contribute to the improvement and further development of the courses in question.

The ISI experiment, which concerned the recruitment of school-leavers in the area of computerized administration, took place from 1985 until 1990. The aim of the project was to set up and implement practical courses for school-leavers and the young unemployed at lower and medium levels of computer administration. ISI provided courses for two functions, those of Micro-computer Operator and Applications Programmer. Regional foundations were set up to support the implementation structure of the ISI courses. The main task of these foundations was to recruit and select course participants, to find lecture halls for the theoretical studies in regional schools or vocational training schools, as well as training places in companies and other institutions. The national coordination of the ISI project was in the hands of the ISI project management committee formed by CIBB and ECABO.

On the basis of the ISI experiences and research into the needs of trade and industry in this area, ECABO decided to create four new information-science courses as regular courses of study in apprentice-system education. The old course in Micro-computer Operator Studies was split up into Application User Studies, Application Management Studies and Application Development Studies.

1039

The course in Application Programming was restructured. The national coordination of the new courses in information science rests with ECABO.
Two additional changes that took place after the ISI experiment should be mentioned. First, the regional foundations created by ISI were converted into JCA foundations (Joint Courses Activities foundations). The latter are regional cooperative groups of companies with the aim of making vocational training less dependent on the often restricted and economically sensitive efforts of individual employers. The second change was that JCA foundations began to actively train course participants during the practical part of basic training, in a specially developed protected working environment (PWE). As simulation environments or practical training centres the PWE's form a link between the acquisition of basic qualifications at the beginning of a course and professional practice or further education in the required professional qualifications. PWE's can be regarded as intermediate provisions: educational provisions made for the transit zone between regular schooling and full participation in the labour system. In such a transit zone one can prepare for and adapt to quick changes in the need of industry and trade in a flexible way.
We refer to evaluation research into the ISI courses as the ISI research, and to research into the new information-science courses as the NIC research. The research questions were identical for both projects and are as follows:

- To what extent were the aims of the courses realized?
- What factors were influential in the realization of these aims?
- What recommendations for improvement can be given?

Choice of variables and research method

The choice of variables in both the ISI research and the NIC research was based on the following:

- The goals of respectively the national ISI project and the new information science courses of ECABO.
- Taking into account the wishes of CIBB and ECABO which have commissioned the research projects.
- Relevant literature on evaluation research (Cronbach 1982; Glass & Ellet 1980; Van de Grift 1984; De Groot 1986; Hofstee 1980; House 1977; Stokking 1982, 1984; Stuffle-beam et al 1972) and implementation of innovations (Fullan 1982; Lagerweij 1985).

In order to find answers to the research questions data for both projects were collected by means of questionnaires and interviews from coordinators of the regional foundations and JCA foundations, as well as teachers, company counsellors and the participants themselves. Additional data for the ISI research were obtained from former students, and from companies. Also ECABO examination data and CIBB employment data from former students were available. Additional data for the NIC research were obtained from vocational training school coordinators, PWE supervisors and ECABO counsellors.

Results

In the context of the three central research questions the results of the ISI and NIC research were as follows. The aims of the courses were achieved to a reasonable or satisfactory degree. Some remarks are in order in this connection. In the ISI

courses the "quality" of the students turned out to be and important factor in course examination results and the post-educational employment situation. The level of previous training and the degree to which the planned course had been thought through by the students correlate positively with examination results. The level of previous training and not being unemployed at the beginning of the course correlate positively with the post-educational employment situation. This points to the importance of the procedures and criteria used in the recruitment and selection of course participants. Another significant factor in the students' results was the degree to which materials were used.
An important factor that influenced the results of the new information-science courses are learner characteristics as there are a relatively high age and knowledge of computers at the start of the course. In contrast to the ISI research, the level of previous training is not a predictor of examination results in the NIC research.
In the ISI courses also the quality of the teachers, training places and supervision turned out to be important factors in course examination results and the post-educational employment situation.
Other important factors that influenced the results of the new information-science courses lie in the area of transfer of knowledge and supervision, links in substance and form between course components, the amount of work carried out in the training place, the degree to which practice manuals are used, and what is done during the regular theory sessions at the vocational training schools.
Results were compared with certain trends found in the literature on vocational training.
Both in the ISI research and the NIC research recommendations were made to improve the apprentice-system courses in information science. Most recommendations in the ISI research were used to develop the four new information science courses as coordinated by ECABO. Some recommendations in the NIC research are the same as in the ISI research. These concern needs assessment research on information science qualifications in organizations, attention to a directed recruitment and selection of students, and guiding the concrete activities of the students in the PWE, the training places in companies and during the theoretical studies in regional schools. In the meantime, ECABO exerts oneself to realise these recommendations.
Other recommendations that were formulated in the NIC research are to meet the wants of those concerned for having more influence in the development of the courses, and for having more contact with the regional companies. Also more attention should be given to the improvement of the mutual relations and to the improvement of the tuning and cooperation between the different parties involved in the development of these courses. More ECABO counsellors have to be appointed, and research has to be carried out for a further improvement of the courses, and for the careers of ex-students. Finally, attention has to be paid to the improvement of the organizational structure of the information science courses. These recommendations too are being worked on by ECABO.

Discussion
From a comparison between the two evaluation research projects a performance profile of the courses is distilled. This profile constitutes the setup and realisation of

1041

the present courses in information science within the apprentice system related to insights gained from the present evaluation research. Based on this performance profile and outcomes in literature concerning developments in vocational education and practical training (COLO et al 1980; Frietman 1990; Van Hoof 1987; Kanselaar et al 1987; PCBB 1979; Schellekens 1980; Stokking 1988; Treep & Pieters 1988, 1989; De Vries & Van Wel-Berretty 1986) we introduce an alternative setup of apprentice-system courses in information science.
In the present setup of information science courses the PWE and the regional vocational school play an important role. To achieve a better tuning of activities which are now organized separately in PWE's and vocational training schools, these institutions should be transformed to one organization for both theoretical and practical training and application.
This also means that the positions of PWE supervisors and teachers will be no longer separate, but have to be integrated into one new position, namely that of a "practice-teacher", who must have an allround expertise.
The role of the ECABO counsellor is also very important; he or she is the key figure between the students, the training centre and the companies that offer training places. The ECABO counsellor should therefore have more authority in controlling the tasks as executed by the "practice-teachers" and the company counsellors.

References

Boekenoogen, M.S., Eberg, J.W., & Stokking, K.M. (1989). Evaluatieonderzoek Informatiseringsopleidingen ECABO. Interimverslag 1. Utrecht: ISOR.

Boekenoogen, M.S., & Eberg, J.W. (1990). Evaluatieonderzoek Informatiseringsopleidingen ECABO. Interimverslag 2. Utrecht: ISOR.

Boekenoogen, M.S., Eberg, J.W., & Stokking, K.M. (1991). Opleiden door samenwerken. Uitvoering en resultaten van leerlingwezenopleidingen voor informatiseringsberoepen. Deel 1: Eindrapport: deel 2: Uitgebreide beschrijving van de uitvoering van de opleidingen in het cursusjaar 1989-1990. Utrecht: ISOR.

Stokking, K.M. (1988). Het gebruik van tests als selectiemiddel voor toelating tot automatiseringsopleidingen: een proefonderzoek in het kader van het ISI-project. Utrecht: ISOR.

Stokking, K.M., & Eberg, J.W. (1989). Opzet, uitvoering en resultaten van informaticaopleidingen in het kader van het leerlingwezen. Eindrapport van het evaluatieonderzoek van het landelijke experiment "Instroomproject Schoolverlaters Informatiseringsberoepen" (ISI). Utrecht: ISOR.

Stokking, K.M. (1990). Evaluatieonderzoek Informatiseringsopleidingen ECABO: aanvullend verslag over de startfase. Utrecht: ISOR.

COLO/GSB/PCBB/BEVAM/ECABO/SOM (1980). Mantelproject beroepen- en functie-analyse. 'Een oriëntatie op de afstemming beroepsonderwijs-beroepspraktijk'. Den Haag.

Creemers, B.P.M. (1984). Kwaliteit van onderwijs. Wolters-Noordhoff,Groningen

Fullan, M. (1982). The meaning of educational change. New York/Toronto, Teacher College Press/OISE Press.

Grift, W. van de (1987). De rol van de schoolleider bij onderwijsvernieuwingen. VUGA, Den Haag.

1042

Grift, W. van de, & Lagerweij, N.A.J. (ed.) (1988). Hoe verbeteren we onderwijs? ABC, De Lier.
Hoof, J.J. van (1987). Arbeidsmarkt als arena. Proefschrift., Amsterdam.
Kanselaar G. et al (1987). MTO-uitstroom-project. 'Onderzoek naar de aansluiting van het MTS-onderwijs naar de arbeidssituatie'. VOU/ISOR, Utrecht.
Lagerweij, N.A.J. (1985). Theorie van onderwijsvernieuwing. In: Onderwijs: bestel en beleid, deel III: Ontwikkelingen. Wolters-Noordhoff, Groningen.
PCBB (1979). 'Begeleiding bij het leren in de praktijk'. Interne notitie door K. van den Heuvel. Den Bosch.
PCBB (1979). 'Leren in de buitenschoolse leersituaties: een schets'. Interne notitie door Th. Ellerman. Den Bosch.
Scheerens, J., & Creemers, B.P.M. (1989): 'Conceptualizing school effectiveness'. International Journal of Educational Research, 13, 691-707.
Schellekens, H. (1980). 'Praktijkleren. Een andere kijk op leren in praktijksituaties'. PCBB. Educaboek, Culemborg.
Schermer, A.F.K. (1991). Implementatie van een onderwijsvernieuwing. Case veldwerk. APS, Amsterdam.
Stokking, K.M. (1988a). 'Beroepsopleidingen tussen vraag en aanbod. Onderwijskundige aspecten'. ISOR, Utrecht.
Stokking, K.M. (1988b). 'Het gebruik van tests als selectiemiddel voor toelating tot automatiseringsopleidingen: een proefonderzoek in het kader van het ISI-project'. ISOR, Utrecht.
Vries, B. de, & Wel-Berretty, J. van (1986). 'Tien jaar buitenschoolsleren'. ITS, Nijmegen.

HOW SHOULD RESEARCH ON VOCATIONAL AND PROFESSIONAL EDUCATION REACT TO NEW CHALLENGES IN LIFE AND IN THE WORKSITE?

Frank Achtenhagen, Georg-August-Universität, Göttingen, Germany

- New challenges for vocational and professional education can be mainly characterized by the following trends:
  (a) Demographic change: lacks within the population pyramid; growing heterogenity of age; mix of nationalities; differences of prior knowledge. (b) Increasing internationality of commerce. (c) Systemic rationalization within organizational structure. - These trends lead to more complex tasks and problems in the worksite. This objective needs occur together with general problems of the youth which provoke a change of values and attitudes: loss of belief in authority; stimulation overflow; fear of ecological disaster; prosperity-influenced behaviour; etc.
- Research on vocational and professional education has to react to these challenges: The youth must be prepared to actively and consciously fulfill the tasks on the worksite; but education also must develop and stabilize the personality. Research, therefore, has to thoroughly describe and explain the tasks and problems and to analyze and to construct teaching-learning processes in a new way.
  The central message is: We have to change the focus of our thoughts on vocational and professional education: away from organizational and institutional patterns towards process structure.
- At the moment, neither vocational schools nor enterprises are prepared to adequately fit its procedures and arrangements to these new challenges in the worksite and in real life. Two reasons have to be mentioned: (a) the very traditional content structure, (b) the missing flexibility of teacher behaviour.
- The actual content dimension in (German) vocational and professional education, e.g. in schools and enterprises, can be described as followed (with a focus on commercial education); The content units are linearized, chopped into pieces, in distance to economic and commercial needs, and they are wrongly mixed with an overrepresentation of very 'traditional' units (as law). The actual demand of enterprises (and also: of real life) cannot be satisfied by that content and curriculum structure. The students and apprentices also have problems with that structure: The organization of the content units and the corresponding linear

1044

teaching procedures very often lead to patterns of rote learning; the mindfulness of the goals and content is kept hidden from the students.

- Teachers, trainers and students are costumed to stay with these lists of goals and content units. Normally, they are not able to effectively handle more complex teaching-learning problems. We find, therefore, a great gap between wonderful proposals of simulation games, case studies, simulated offices, projects, etc. and their actual use in the classrooms.
- The research-based development of new forms of effective teaching and learning has, therefore, to be especially aware of at least two crucial problems: (a) One has to think on the content dimensions of the whole curriculum. (b) One has to think on the mid-term and long-term dimensions of the development of the students' personality and their higher order abilities (as problem solving or so-called "keyqualifications"). This point, especially, marks the bridge to a consideration of learning with regard to the new challenges and life-related learning. The exploration of the time perspective within the years of vocational education may give decisive hints for an effective lifelong learning which is more than refreshing of a knowledge of facts.
- The research should be organized as a combination of curriculum research and teaching-learning research; by than, it is focussed on perspectives of time and action. This research leads to new complex teaching-learning arrangements, e.g. computer-based simulation games, case studies, workinganalogue learning tasks, learning environments etc. The construction of these arrangements should satisfy at least five criteria; (a) The goals and the content of each arrangement should be of importance for the worksite. (b) The goals and the content should correspond to modern scientific research for the specific field of practice. (c) the students and the apprentices should judge the goals and the content as subjectively important and mindful. (d) the goals and the content must be conceivable and understandable. (e) The teachers and trainers must accept the arrangement and its background-theories and they must be thoroughly trained to effectively handle it.
- Short-term and mid-term evaluation procedures for these teaching-learning arrangements are related to knowledge of facts, problem-solving with regard to tasks in the worksite or simulations of the worksite, and 'key qualifications'. The concurrence of cognitive, motivational and emotional factors has to be controlled. These procedures will be illustrated by data.
- The long term evaluation has to be run in the enterprises and in the field of recurrent education.

LABOUR MARKET AND OPPORTUNITIES OF BECOMING SKILLED IN ENGLAND AND THE FEDERAL REPUBLIC OF GERMANY

Martina Behrens, University of Bielefeld, Germany, & Alan Brown, University of Surrey, United Kingdom

Recruitment and the labour market

In England a varity of routes into skilled employment exists, involving mixtures of employment, education and training. In the FRG the vocational training is dominated by the more or less compulsory dual-system. Indeed, 60-70% of young people - virtually all those not going to university are in an apprenticeship, working for their employers, attending the Berufsschule and special training courses. Thus, while the German system produces on the whole a surplus of skilled labour - the English system at best tends to be tailored to immediate requirements. Central issues of this paper, therefore, are the influence of labour market condition and changing recruitment practices on local training opportunities for school-leavers. In the German dual system systemic factors are most influencial in the shaping and patterning of local training provisons. Whereas in England there appears to be a direct relationship between provision and the state of the labour market. In this context the paper examines the strategies of expanding and contracting training opportunities in more detail.

Method

This paper draws upon a study of vocational preparation in contrasting labour markets in England and the Federal Republic of Germany (Bynner/Roberts (eds) Youth and Work: transition to employment in England and Germany. 1991). The research compared the experiences and attitudes of 160 16 - 19 year olds in each of four towns. The young people were matched in terms of route and destination with those from a town twinned on the basis of the state of the labour market: Liverpool and Bremen (contracting); Swindon and Paderborn (expanding). In addition, key personnel (company trainers, staff from colleges, vocational training bodies, labour administration) were interviewed to examine the education and training curriculum and context, policies, and labour market opportunities. Further contextual information on governmental policies were gained through official papers and statistics.

1046

Results

The high demand (and intense competition among employers) for young skilled workers in Swindon meant that youngsters could be confident that they would be able to enter employment in their chosen sphere. The labour market opportunities were such that a much fuller range of avenues (including jobs) were open, and YTS never fully established itself as a popular alternative, except in the cases where it was part of, or at least opened up the possibility of, a full apprenticeship, similar training or employment with pros- pects. In such cases, particularly if their status was `employed YTS' and the scheme was based upon a single employer, they may have been aware that administratively they were on YTS, but they saw themselves primarily as `employed'. Some large employers used YTS to supplement their recruitment, whereby they took on youngs- ters who did not necessarily meet their `normal' entry requirements but who they would be prepared to keep on if they made satisfactory progress. Examples of this practice could be seen across occupational areas: for example, with insurance accounts clerks as well as engineering technicians. The success of some of those who previously would not have been deemed `suitable' had the effect of causing some of the companies to review and `open up' their recruitment processes. One major local employer acknowledged that those recruited onto YTS, with few if any educational qualifications, sometimes compensated for this with keeness and enthusiasm. This had taught the company that to take one isolated marker of ability (educational attainment) could be misleading: "there is now a preference to select upon a range of qualities: that is, to weigh up other factors and all categories of information".

More generally within YTS there were some schemes, which while fulfilling the minimum criteria appropriate at the time, in practice gave only limited education and training, the rest of the time being spent in (possibly not very demanding) work experience. The latter rarely occurred in Swindon, and although there were some schemes between the two extremes in Swindon these were getting progressively `squeezed', particularly by the attractions of permanent employment, but also by other routes which offered more substantial education and training.

The high level of youth and adult unemployment in Liverpool meant that YTS was still the most likely alternative to education, rather than employment. In Liverpool, not only is there a far more limited range of employment opportunities, but given the greater supply of youngsters, Liverpool employers are much more likely to use firm-based YTS as an `extended probation' and not have to make selection decisions immediatetly. Therefore, youngsters in Liverpool were less likely to get `employed status' YTS for comparable positions.

In Paderborn and Bremen the high demand for apprenticeships over the last ten years made the employers choose their apprentices very carefully. Standing entry require- ments for certain professions were ignored to the extent that employers asked for the highest possible school-leaving certificate (Abitur) wherever possible. This led to a devaluation of the other certificates. In the early nineties the situation changed. The demographic decline, the falling numbers of school-leavers and rising numbers of university entrants made employers compete for apprentices. In some crafts espe- cially in construction it is found very difficult to recruit apprentices. In Paderborn employers complain of a decline of applications in clerical areas of about 50% and in craft areas of about 70%. As a result young

people now expect higher allowances in some craft apprenticeships (e.g. in the building industry). So far this development has not had any significant effect on the number of applications. However, companies in all vocational fields have had to reconsider their recruitment practices and return to acknowledging all school-leaving certificates. This is true for the prosperous labour market in Paderborn whereas the declining labour market in Bremen has reached an equilibrium between supply and demand for apprenticeships for the time being.

1048

ASSESSMENT OF THE EFFECTIVENESS OF CAREERS COUNSELLORS IN SECONDARY EDUCATION

E. Voncken, E.C. Mellink, Y. Bosma, & P.N. Karstanje, Centre for Educational Research (SCO), University of Amsterdam, Amsterdam, The Netherlands

Introduction

In the beginning of this century problems in choosing jobs were thought of as problems resulting form lack of knowledge. The answer to this problem was not very complicated: providing pupils with information. Since this time, however, many changes have occurred. In the Netherlands nearly all schools for secondary education have one or more teachers operating as careers counsellors. These careers counsellors are concerned with offering educational and vocational guidance; informing and helping pupils in making a choice for further education and/or a profession. In many of these schools the counsellors have a specific function, -distinguished from the tutor's or mentor's tasks-, whose activities vary from providing guidance on the learning process to assistance with personal development/problems.

Recently much attention has been paid to educational and vocational guidance. Judging from several reports and notes by the government, the expectations for positive results are running high. Educational and vocational guidance is mentioned in the same breath as in emancipatory policies, raising educational outcomes, changing stereotypical choices (e.g. `Choose Science'), transition from education to work or to further education, preventing early school leaving/drop-out, internationalization, etc.

Throughout the years, educational and vocational guidance has been developed from a (single) advice linked to the incidental choice (e.g. of examination subjects) to a process-based guidance-system (Wijers 1987, Verijdt en Diederen 1987). This implies a shift to a career policy and therefore the involvement and cooperation of several persons, within or outside the school.

Problem and research method

In the past years both research and theory have focused primarily on the structure of vocational and educational guidance and on the process of development of (vocational) choices. On the issue of vocational choice several theories descending from different disciplines have been formulated. For example, the psychological approach underlines the importance and influence of internal

determinants for vocational aspirations (e.g. abilities, standards, ambitions). In the economical perspective the dynamics of the labourmarket define the choice for occupations. Besides the two approaches mentioned, the sociological perspective emphasizes on external or environmental influences (Van der Aa e.a., 1991). According to Wiegersma (1990), however, the choice of a job is, like many other events in life, coincidental. In recent years consensus on a synthesis of these approaches has grown, from the perception that each of them definitely represents part of reality (Wiegersma, 1990).

Less attention has been paid, however, to the effectiveness of careers teachers. In this paper we will highlight the problem of determining effects of careers counsellors' activities. There doesn't seem to exist agreement on the criteria of effectiveness of career guidance. In the research project on `Performance and effectiveness of careers counsellors' a partial solution has been conducted by asking many actors for their perception of the relative influence of the careers counsellors' actions. This way of acting causes some methodological problems that will be illustrated in the paper.

From 1990 to spring 1992 the Centre for Educational Research (SCO) carried out an investigation into the performance and effectiveness of careers teachers in secondary education. In this research project not only the activities of careers counsellors are studied; also the effectiveness of their work performance is a topic of special attention.

In short, the study consisted of four parts:

- A national survey on the policy of schools, the organisational structure, the content of the educational and vocational guidance provided at the schools and the role and position of the careers counsellor, amongst 200 school principals and 200 careers teachers;
- An in-depth study of twenty-five schools. To gain an better understanding of the content and structure of the educational and vocational guidance information was provided by principals, careers counsellors, pupils, parents, mentors and teachers;
- A panel study with one panel of principals, whereas the other panel was composed of policy makers and representatives of national institutions and organisations;
- Six case studies, in which the guidance and counselling system per school is described more coherent.

Methodological problems

The measurement of the effectiveness of the vocational and educational guidance causes some difficulties, related to:

- the purpose of vocational and educational guidance;
- (lack of) standards for a right choice/development of choice;
- lack of) development/formulation of theories on the process of choice;
- the careers teachers' role and their relative contribution to this process.

As we define the most important purpose of vocational and educational guidance to help pupils in making a `right' choice, the first difficulty we are faced with is the definition of `right'. In defining this we meet with different interests of several 'parties' concerned. For example: a pupil may choose according to his/her interest and abilities, but this may conflict with the social demand for emancipatory choices, with the availability/accessibility of certain jobs, or with certain policies of

the school. From the employment situation's point of view a right choice may be a match with current job opportunities. From schoolpolicy's point of view a right choice can be seen in terms of financial outcome, but also the resulting student-enrolment. It may not be in the school's best interest if pupils switch to another schooltype before having passed the final examination or choose a subject in such a way that it causes problems in the timetabling.
Moreover, if the educational or vocational choice is not valued at its result, (i.e. the content of the choice), but at the process of development, (i.e. the skills needed to make desicions), changes in the pupils' choices may reflect a well-considered decision as a result of a learning process. These changes therefore should not necessarily be seen as a bad choice (cf. Gottfredson, 1981). An additional problem is the implicity of the guidance; most careers counsellors do not give explicit, directive (standard)advices, as a consequence of which pupils do not trace their choices to the counsellors activities.
As we have seen the choices pupils make are determined by a large number of factors of which only a small part can be measured. Moreover, the pupil's reconstruction of important influences tend to be distorted (`self-serving attribution bias') (Weerdenburg, 1987).
Several studies reveal the influence of socio-economic background. Some of the results show the implicit reproductive role the educational and vocational guidance serves especially for middle class students.
It may be clear that one simple criterion/standard by which the effect of guidance can be measured does not exist. It would be too simplifying to judge the effect by its outcome, meaning the destinations of school leavers. However, a student monitoring system including this information could serve the careers teachers in providing them with a clear and systematical view on patterns in destinations. This information could be used as a basis for policy plans or steps.
Given the influences described above, the scope for a careers counsellor to perform actions is limited. The difficulty is therefore to determine the *relative influence* of the careers counsellors' activities.
Considering these restrictions, we have chosen to include six groups of respondents in our study: the schools' principals, careers counsellors, pupils, their parents, mentors and teachers.
One possible indication for the effectiveness is the appreciation of the parties concerned for the activities of careers counsellors. We tried to operationalize this by relating the appreciation to the interests of the parties. For example: school principals were asked for their opinion on the effectiveness of the careers teacher in terms of preventing `wrong' choices, enrolment and financial effects.
As for the students and their parents we inquired after actual activities in vocational and educational guidance, the perceived use and the influence on the process of choice (in terms of `*I've started thinking/talking more about the future', I regret my choice of a certain subject* etc. Furthermore students and their mentors/teachers were asked to point out the school year in which students profit most from educational and vocational guidance. And the careers teachers for instance filled in their needs for further training.

Results

The results indicate that the principals represented in this study are pleased with their careers teachers and their role in the educational and vocational guidance. The answers to the questionnaire showed slightly some of the earlier mentioned bottlenecks: for example the school principals and careers teachers agreed on the statement that `*many wrong choices of examination subject are prevented through the guidance of careers teachers'*.

In the interviews, however, we found more support for our belief that career teachers have to reckon with other school interests (timetables, free periods, etc). Almost all principals share the opinion that the school could hardly do without careers counsellors.

A second indication of the effectiveness is found in the extent to which students consult the careers counsellor. It was found that students primarily discuss matters of choices with their immediate environment (parents, friends). Within the school they consult the careers teachers, more than they consult other teachers/tutors. This tendency might not be exclusive, but complementary: in case of doubt or problems the careers teacher is asked for help. Boys and girls do not differ as to the extent they consult careers counsellor, although they do differ in appreciation: girls value this contact higher than boys do.

It looks like boys use their consultation to consolidate their choice, whereas girls use this contact more for orientation.

As for the best way a careers counsellor could assist in problems most students prefer his/her help by providing information and letting them free to choose on their own. One-third of the students prefers help with reflecting on possibilities and consequences.

The most important effect of the vocational and educational guidance on students is their increased concern with and reflection on the future. Secondly students mention an increase in self-knowledge, in their view of their abilities, interest and personality. Decision-making skills however are seldom learned.

In order to be able to review conditions for a well functioning system of educational and vocational guidance the careers teachers were asked for their needs for further training. In addition demands and conditions realized by school heads and deputies were listed.

It was found that nearly all careers teachers want to participate in additional training, particularly on newly developed methods for educational and vocational guidance.

Finally we found a positive relation between the school policy, the structural position of the educational and vocational guidance in the curriculum and the number of facilities available for careers teachers.

References

Aa, R. van der, Hofman, W.H.A., & Vlaanderen, A. (1991). Determinanten van beroepskeuze in verschuivend perspectief, in: Onderwijs en Samenleving, red. J. Dronkers, G.W. Meijnen en H. Dekkers, ORD 1991, SCO, Amsterdam.

Gottfredson, L.S., (1981). Circumscription and compromise: a developmental theory of occupational aspirations, in: Journal of Counselling Psychology Monograph, 28, number 6.

Karstanje, P.N., Mellink, E.C., Voncken, E., & Bosma, K.Y.G.C. (1991.). Performance and effectiveness of careers counsellors. Interim report, SCO, Amsterdam.

Verijdt, H., & Diederen, J. (1987). Determinanten van beroepskeuze, OSA-werkdocument W39, Den Haag.

Wiegersma, S. (1990). De geschiedenis van het beroepskeuzewerk: een gespleten ontwikkeling. In Intermediair, 26e jrg., 16 februari.

Wijers, G.A. (1987). Een ontwikkelingspsychologisch concept van beroepskeuze en arbeidsmarktprognoses, OSA-werkdocument nr W 40, Den Haag.

RESPONSIVENESS AND ORGANISATIONAL EFFECTIVENESS IN VOCATIONAL EDUCATION AND TRAINING

Jan N. Streumer, Department of Education, University of Twente, Enschede, The Netherlands

Presentations/participants

- The development of vocational education and training: industry and service compared;
  E. de Bruijn, SCO, University of Amsterdam, & A.F.M. Nieuwenhuis, RISBO, Erasmus University Rotterdam, The Netherlands.
- Responsiveness of vocational education towards new technologies: CNC–cutting and flexible production automation (FPA);
  J.N. Streumer, & A. Feteris, OCTO, University of Twente, The Netherlands.
- Developing the quality of responsiveness through modules;
  S. Clark, Scotland, United Kingdom.
- Responsiveness of vocational education and training in the printing industry;
  J. v.d. Berg, & B. Hövels, ITS, University of Nijmegen, The Netherlands.

The occasion of organising this symposium lies in the Dutch part of the study "The changing role of vocational and technical education and training (VOTEC)" endorsed by the Organisation of Econmic Co–operation and Development (OECD).
Four Dutch research institutes (ITS, SCO, RISBO and OCTO) were invited by the Ministry of Education and Science, Directorate for Vocational and Adult Education, to contribute to the Dutch country report.
The VOTEC–study elaborates on I. the historical context of vocational education and training; II. recent developments in selected occupational and training areas; changing objectives, conditions and policies; III. ensuring the responsiveness and effectiveness in vocational training and education.
The afore–mentioned four research institutes were each, besides a contribution to part III of the OECD/VOTEC–study, responsible for an illustrative case study in which the way vocational education and training were responding to changing needs and conditions within a branch/sector of industry and/or business was examined.
The following branches/sectors of industry/business were chosen: tourism, printing industry and installation engineering. Besides these three branches/sectors of

1054

industry and/or business one case study was aimed at the way vocational education and training were reacting to structural technological innovations; in particular CNC–cutting and flexible production automation (FPA). In the symposium also a contribution from the UK is included. It is a contribution from dr. Sheila Clark who was responsible for the Scottish part of the OECD/VOTEC–study. She will report on the way the Scottish vocational further education and training policies have sought to encourage access, retention, and return; improve links between stages of education and training and between vocational further education and industry; reduce academic–vocational divide; and encourage quality development. Dr. Clark will illustrate these actions on the basis of a case–study.

SCOTTISH OUTCOMES, WITH REFEFERENCE TO CONSTRUCTION AND TOURISM

Sheila Clarke, Independent researcher and developer, Bearsden, Glasgow, United Kingdom

The broad picture of VOTEC in Scotland is painted in David Raffe's paper to this conference . He diagnoses the climate of uncertainty and change for the economy, society, government, and individuals in which VOTEC must operate. He overviews modular responses to these pressures and issues of increasing flexibility through modular provision. The purpose of this paper, based on an OECD study , is to highlight recent initiatives in Scottish VOTEC with illustrations from the construction and tourism industries.

Changes in construction and tourism

Both the construction and tourism industries have undertaken self critical evaluations which have influenced VOTEC reforms They both report a degree of low public image, eg. regarding quality of buildings and of tourism services; the efficiency of work forces and managements; instability of employment; and lack of progression prospects, poor status, pay, and working conditions for employees. The building industry has a tradition of VOTEC based on single-craft apprenticeships starting at age 16. Changes in practice in the industry are creating gaps between traditional skills and those needed in a modern industry. These relate to new materials and techniques, increasing mechanisation, and use of information technology. Trade boundaries are breaking down and changes in work practice place a premium on communication, management skills, team working, and individual autonomy. The tourism industry (catering, accommodation, travel, leisure activities and visitor attractions) has a history of training provision, but on the whole has been characterised by a low proportion of trained workers and a lack of qualifications in some specialist areas, and, as in construction, a lack of broad based qualifications at non advanced levels. The urgent need for skilled and flexible personnel is making established patterns of provision look cumbersome. In tourism in particular, small businesses report that much training available is not relevant to their needs. Training programmes have followed rigid patterns, been designed largely for the young, and have been difficult to access

other than by standard routes. Training providers have perceived business as being apathetic about training and so have provided what they can, not what is needed .
It is against this background that construction and tourism VOTEC (along with other forms of VOTEC) are in a state of redefinition and development. There is growth of partnerships between industry and VOTEC which are improving the content and relevance of training and associated qualifications; improved institutional arrangements which are making training opportunities more accessible to young and old alike; and quality improvements which are stimulating a demand for training and improving the status of vocational education.

Improving partnerships between education, training, and the world of work
Through the 1970-80s there has been a relatively remote relationship between VOTEC and the world of work, with VOTEC only one option in keeping ahead of the competition, together with business practices such as change or rationalisation of product or service, automation, or franchising. Despite predictions of a declining number of young entrants to the labour market, employers weigh up costs of recruiting or poaching staff with needed skills against the costs of in-house staff development. Tourism employers may not see business benefits in investing in employees, who are often part-time, temporary, and viewed as low calibre. Evidence from the engineering sector implies that firms can often sustain output of sophisticated products, in spite of a largely unqualified work force, by optimising qualified labour between tasks and by relying on uncertified skills acquired through experience. A commitment to the training of the work force has not been seen as a normal long term business strategy. In general in the UK,VOTEC has been seen as primarily a national and local government concern.
A radical response has been change in strategic organisation of VOTEC with considerable devolution of decision making. Youth and adult training schemes are, since 1991, devolved from the UK-wide Training Agency to Scottish Enterprise agencies and local enterprise companies. These have the dual role of promoting the development of industry and commerce and encouraging and organising training relevant to the local needs and problems of Scottish industries. The arrangements should, once the system is embedded, assist in the provision of accurate advice on industry's qualification requirements and on the types and quality of VOTEC needed. Scottish Enterprise intends to shift the balance of support from youth and adult unemployed training to programmes aimed at upgrading skills. Finance for much youth training will be funnelled through a training credit scheme for all young people leaving full time education at 16 or 17. In the first place training credits will be particularly targeted at small and medium sized businesses in sectors such as tourism with previously poor training records. An historic change in the management of further education will take effect from April 1993 when the Scottish Office Education Department (SOED) takes over the duty of providing most further education from the local authorities. Responsibility will be devolved to further education colleges which will be corporate bodies managed by Boards, at least half of whom will be employers. The reforms will, the government hopes, help to support colleges in efforts to become more responsive to business customers since these same customers will play a major role in planning and operation of colleges.

1056

BEST COPY AVAILABLE

Meanwhile, much of the VOTEC curriculum is in the process of revision, involving industry lead bodies (UK-wide government sponsored bodies representative of employment interests). They are developing occupational competences and assessment procedures which are incorporated into vocational qualifications offered by an awarding body. In Scotland, the Scottish Vocational Education Council, SCOTVEC, offers Scottish Vocational Qualifications (SVQs). This key role for industry is a significant shift in balance of power over vocational qualifications. SCOTVEC has always involved employers in advising on the shape of syllabuses and assessments, including National Certificate modules, but, even in industries with a good record of employer involvement such as construction, educationalists have played the major role. Their concern has tended to be with the quality of teaching input and the quality of learning resulting from this input. Lead bodies develop standards by functional analysis which identifies essential outcomes of learning in the form of observable competences which are key features of performance at work.

The proposals described above provide potential for coherent policy control as VOTEC becomes the concern of a new partnership between SOED, SCOTVEC, industrial and business concerns led by enterprise agencies, and training providers. The question is, can SVQs be implemented as planned? The UK standards development programme which led to their design has been an ambitious development which aims to ensure that eventually the majority of the work force are certificated as competent in the job role which they occupy. SVQs recognise that attainment of standards in an educational environment is not automatically seen as competence in the workplace. However, a great deal needs to be done to encourage employers to take up the new qualifications. Industry are inclined to develop people at work but may not seek formal recognition in terms of accredited work place learning. The issue is not merely one of recognition. The delivery of SVQs is in many cases impossible without strong employer contributions. Can employers develop the training and assessment infrastructure which will be necessary?

Another issue of SVQ implementation may have to be addressed in the near future. There are some indications that employers (and their staff) may develop an enthusiasm for work based and competence based programmes of education and training and they may also see the benefits of certification. However, they may wish the competences to reflect the needs of organisational development (or perhaps individual development) rather than being exclusively tied to national standards. In other words, the SVQ approach may become integral to quality management strategies, but only if the system is flexible enough to take account of the needs of individual organisations. If SVQs become an integral part of the quality revolution this may ensure their success.

Improving links between types and stages of education and training

The esteem in which VOTEC is held has improved because of recent reforms in the National Certificate and advanced VOTEC and better articulation between the two. Vocational education, whether taken full time in colleges, in training schemes, or release from work, now provides routes to higher education, with ladders being built even in areas such as tourism which have previously had little scope for progression. Prospects for parity of esteem are helped by the ending of the binary

line in higher education. This will no longer be divided into two separate spheres, one academic, one vocational and therefore all types of degree will be in one system and differences in esteem greatly reduced. This could have a backwash effect in enhancing the status of non advanced vocational education.

The vocational ladder has a weak base compared to the academic ladder since modules in schools have not generally provided coherent qualifications. There is a growing belief that general education opportunities and core skill development should be accessible to all young people so as to encourage a common culture, an increased awareness of society and economics, and an ability to play a useful and personally enriching role in society. Such a social context is important for those occupying high level occupations and the fact that general education and core skill development is not, by and large, pursued in vocational programmes forms a barrier to equality of opportunity. Improvements are expected through proposals for new upper secondary school qualifications. Highers would be replaced by academically-oriented Scottish Baccalaureates and vocationally-oriented Scottish Certificates . Employers and higher education should benefit from the new arrangements. They will be presented with more coherent and relevant qualifications on which to base recruitment. They will recruit more generally educated young people who have been trained in core skills and carry accreditation of their achievements in these skills. Such a change would begin to address the gap between VOTEC practice and employment needs in construction and tourism since young people would be better prepared for multi-skill training and future mobility.

Vocational group awards will all be unit based, building on National Certificate modules and parallel developments in SCOTVEC's Advanced Course provision. Good credit transfer arrangements are proposed between the new upper secondary awards and, if current trends in higher education towards modularising their provision gather pace, we are close to having the strategy for a single system of credit transfer embracing the whole of post-16 education and every type of education and training centre.

In principle, such unification allows for coordination of provision across the variety of institutions and initiatives, improving access to training and qualifications. The strategy should imply that the range of institutions provide a linked, client-centred service through regional and national networks. Ideally, potential learners would be referred to centres which provide advice on courses, career opportunities, and finance and advice on the most appropriate options and providers to suit learners' individual goals and existing competences. Their needs would be analysed. There would be opportunities for assessment and recognition of prior achievement and individual action planning. Tailored programmes of learning and assessment would be arranged, with place, mode, method, and pace of study to suit individuals (eg.via open learning, accreditation of prior learning, work-based learning, or college or university course based learning). Throughout the selected programmes learners would take responsibility for organising their learning and studying, though there would be support for any learning difficulties and opportunities for additional learning interests to be followed up through flexible, drop in centres, and access to guidance tutors. So far as possible, given constraints of external assessments, assessment would be on demand. Prior to exit there would be another chance for learners to reassess their goals and plan

for progression. This would be a spiralling process, through which most, if not all 16-18 year olds should pass; most would return at periods throughout their lives as there was change in employment and other circumstances.

The strategy just outlined has yet to be implemented in a nationally coherent way. Field tested initiatives have established, for example, the principles for management and delivery of open learning, work place learning, guidance, and assessment and accreditation of prior learning. Through business-education partnerships, novel programmes for individual employers can be developed and delivered quickly. Colleges are adapting work practices and environments in efforts to place students at the centre of the curriculum. Innovative award schemes are emerging. Awards at certificate, diploma, and degree level can be gained by accumulating credit through programmes of work and study; validated evidence of earlier, unaccredited work and study; and evidence of development work carried out on-the job. Tourism, due to support from the enterprise agencies, is an area of recently improved provision in these respects. In general, though Scottish VOTEC students may encounter some of the support described, they are highly unlikely to experience the whole. There is a need for some kind of national initiative which will pull together and consolidate the developments and encourage widespread uptake. A national centre of excellence, for example, could model flexible learning and guidance practice of the highest standards and provide consultancy support for VOTEC centres.

Quality development

An aim of the 1980-9s has been the improvement of quality assurance with development and application of quality standards and associated performance indicators. Incentives for quality improvement are provided through devolution of responsibility for management, curriculum, financial and personnel matters to local levels, resulting in a much more conscious effort to improve quality than there used to be. Moreover, there seems to be a willingness of Scottish agencies to pull together to create a common quality assurance system. However, there is an issue which still has to be firmly addressed. The key resource on which quality development depends is VOTEC staff but the structure of training and qualifications for staff is not tuned to the task. There is a need for a competence based and work based suite of qualifications which adequately reflects the knowledge and skills which staff require for work in a modern VOTEC system. In Total Quality Management terms it may be that the priority is to address the needs of the 'internal customers' – education and training staff. In Scotland a strong case for respite from further policy initiatives is a need for emphasis on support for implementers of policy decisions.

References

Centre for Strategic Studies in Construction (1990). Building Towards 2001. NCG/Building Employers' Federation.

Clarke, S. (1992). Scottish Outcomes: The Changing Role of Vocational Education in Scotland. (OECD pending).

ED. TEED (1992). Changing Skills Needed in the Construction Industry.

Hart, P.E., & Shipman, A. (1991). 'Financing Training in Britain'. National Institute Economic Review, May 1991.

PA Consulting Group (1989).. Training for Advantage. Tourism Training in Scotland: The Way Forward.
Raffe, D. (1992). The New Flexibility in Vocational Education.
SOED (1992). Upper Secondary Education in Scotland. (The Howie Committee Report). HMSO.

THE DEVELOPMENT OF VOCATIONAL EDUCATION AND TRAINING: INDUSTRY AND SERVICE COMPARED

E. de Bruijn, SCO, University of Amsterdam, Amsterdam, & A.F.M. Nieuwenhuis, RISBO, Erasmus University, Rotterdam, The Netherlands

Introduction

In 1990 the OECD initiated a project on the changing role of vocational and technical education in 28 countries. In the Netherlands four sectors of industry and business are analysed on the relationship between economic and technological developments and changes in vocational and technical education.
In this paper we compare the results of the analyses of two of those branches: a rather old branch of industry is the installation engineering (cf. Nieuwenhuis and Steijvers, 1991) whereas tourism and recreation is a rather new branch of the service sector (cf. de Bruijn, 1991).
For the comparative analysis we use the sociological perspective of occupations. Occupations can be seen as coherent structures of tasks and qualifications which professionals use to hedge a specific domain of economic activities (cf. Geurts, 1989). In this perspective occupations are not only seen as a functional category, defined by skills and competences, but also as a strategic category in which the defending of economic markets and positions forms a central element. The development of occupations is not only a proces of defining coherent domains of competences but also a strategic game to conquer own positions on the economic and the labour market. Hedging activities are made concrete with formal and informal ballot procedures, licence systems and so on. Educational qualifications play an important role in this game: they form the formal permission documents. The existance of an educational track towards an occupational domain legitimizes the occupation itself. Therefor education and qualification serve not only the occupational skill formation but also the economic protection of market positions. The development of occupation specific education and training courses is an important tool for policy makers both in a qualificational as in an economic way.
Comparing the two branches involved we focus on the following aspects:

- economic developments;
- economic power and institutionalization; the role af actors;
- strategies and instruments to geer education and labour;
- the role of education in defining occupational domains;
- autonomy of the educational system in designing curricula.

Installation engineering and tourism services

Installation engineering is a rather old branch of industry. The apprenticeship system is the most important way of supplying qualified workers. Organizations of employers and employees control the quality of the training courses. Because of economic and technological changes the traditional division between specialities and jobs seem tohave been blurred. There is even a tendency of integration with other branches of industry such as the building trade and electrical engineering. In the early eighties, the organization of employers initiated analyses of job profiles in order to see to which extent tasks and the required qualifications in installation engineering have been changed. They claimed that the training courses within the apprenticeship system had become obsolete. Thereby they forced changes in the training courses and created an innovative training system.

At the same time the internal qualification routes within the apprenticeship system were modified and external schools for secundary vocational education were asked to develop specific training courses for installation engineering. The policy makers within the branch of installation engineering support this movement in order to consolidate their position as a separate branch of industry and to 'prove' that installation engineering requests specific qualified workers.

Tourism services is a very young business sector which is developing rapidly. The economic activities are very diverse: travel agencies, offices for tourist information, zoo's, campings, museums and so on. In the Netherlands two separate branches exist, each with its own history: the hotel and catering industry at the one hand and the travel, information and recreation business at the other. We will concentrate on the last one, which is also the youngest.

There is (not yet) a clear system of vocational training for this business sector. Professionals in this field differ largely in educational background. In the seventies and eighties, an explosion of different courses occured, mostly private or commercial. At the same time, general schools for secondary vocational education developed specific training courses for the travel, information and recreation trade. This initiative taken by external schools in cooperation with local companies catalysed the structuring of the branch. It forced the policy makers to think about the definition of their expertise regarding the specific sets of economic activities and the specific tasks and required qualifications in the branch. At the same time, it became clear that an own system of vocational training routes would help to stress the unique character of the branch.

Comparative analyses

Occupations can be characterized by two faces: the face of skills and competences and the face of the social and institutional network. The development of those faces is interrelated: the coherence of skills and competences enforces the strength of the intstitutional network and vice versa.

For the installation engineering external developments put pressure on the relation between competences and networks: other industrial sectors try to conquer parts of the economic domain. The installation branch tries to combat these attacks by reinforcing branche specific educational tracks. The tourism branch is actually hedging its own economic domain and in that proces the development of own training courses plays an important role.

So we can conclude that the existance of an own educational system is an important tool for defining and defending economic interests and domains. Within the boundaries of this economic game the educational institutions can play their owninnovative role: the autonomy to designing occupational curricula is rather high in both sectors. The way in which educational institutions handle this challenge is important for the economic survival of occupations and branches.

References

Beck, U., Brater, M., & Daheim, H. (1980). Soziologie der Arbeit und der Berufe. Reinbek bei Hamburg.

Bruijn, E. de (1991). Professionalisering en toerisme en recreatie: beroepsvorming en opleidingen. Casestudie voor het OECD–project "the changing role of vocational and technical education and training". Zoetermeer: O&W/SCO.

Geurts, J. (1989). Van niemandsland naar beroepenstructuur. Een studie over de aansluiting tussen onderwijs en arbeid op het niveau van aankomend vakmanschap. Nijmegen: KUN/ITS (dissertatie).

Krogt, Th. van der (1981). Professionalisering en collectieve macht: een conceptueel kader. Den Haag.

Mok, A.L. (1973). Beroepen in aktie: bijdrage tot een beroepensociologie. Meppel.

Mok, A.L. (1990). In het zweet uws aanschijns.... Inleiding in de arbeidssociologie. Leiden/Antwerpen: Stenfert Kroese.

Nieuwenhuis, A.F.M., & Steijvers, J.R.L. (1991). Responsief beroepsonderwijs voor de installatietechniek. Casestudie voor het OECD–project "the changing role of vocational and technical education and training". Zoetermeer: O&W/RISBO.

RESPONSIVENESS OF VOCATIONAL EDUCATION TO TECHNOLOGICAL INNOVATION: CNC-PRODUCTION AND FLEXIBLE AUTOMATION

J.N. Streumer, & A. Feteris, Center for Applied Research on Education (OCTO), University of Twente, Enschede, The Netherlands

Introduction

The Faculty of Educational Research and Technology (Research Centre for Applied Educational Science) has, at the request of the Ministry of Science and Education, contributed to the OECD study, "The Changing Role of Vocational and Technical Education and Training (Streumer & Feteris, 1991). The goal of this study is "...to identify and compare strategies and policies through which member countries are pursuing objectives, and to examine to which extent and in which ways these strategies and policies are conditioned by social and economic processes, institutional frameworks, and legal dispositions and conventions" (OECD, 1990).

In total, four case-studies were carried out concerning the Dutch situation. Three of these were defined by sector (e.g., tourism, the graphics industry, and installation

engineering). The fourth, which was conducted by the above-mentioned faculty, concerned a study of a specific technological innovation, CNC-production and flexible production automation (FPA).

The development and current position of CNC-production in the Dutch metal industry

Consequences of economic developments

When economic growth stagnates as it did in the sixties, many businesses reevaluate their organisational goals. Since the sixties, businesses have tended to concentrate on clearly defined market segments, and compete on the basis of price and quality. Productivity enhancement, product quality, and the quality of customer service have become important objectives in modern business management. Market strategy is characterized by specialization and standardization.

In the metal industry, firms with mass-production have adopted mechanization and automated production lines (Kayzel, 1986). Because of their client-dependence, firms with customized and small-scale production must enhance their productivity by maintaining flexibility with regard to changing market demand. Scale enlargement occurred through the production of reserve stocks and the introduction of modular design product, through which both the specific desires of the clients can be fulfilled, and the serial production of standard parts can take place.

In the seventies, because of increasing costs, firms switched away from the production of reserve stocks. The most modern production strategies of metal working businesses with customized and small-scale production is directed towards the reduction of stocks, the shortening of production time, and the enhancement of product quality.

Consequences of technological developments

The adaption of production strategies was to a great extent the result of technological developments. Since the beginning of the eighties, production automation on the basis of Computer Numerically Controlled-machines (CNC-machines) has been implemented in various forms within Dutch industry.

The development of CNC-machines went from the original use of conventional machines by skilled workers via the introduction of isolated CNC-machines to the inclusion of CNC-machines in an automated production system. Production techniques change through the integration of automated sub-processes in the production with logistical planning and control systems on the one hand, and computer controlled design and calculation methods on the other hand (Computer Integrated Manufacturing; CIM, Flexible Production Automation; FPA).

Labour market developments

The labour market has been strongly affected by production automation. The initial theoretical projections were very optimistic about the consequences of automation for the quality of labour; later projections were decidedly more pessimistic. The degradation thesis and the polarization thesis played an important role in these pessimistic visions.

Kayzel (1986) concluded that it appears technically as well as economically attainable to make use of automation techniques on the basis of fixed social criteria. This conclusion seems to be becoming even more relevant with regard to the far-reaching new developments in the area of computer technology based techniques. It has become possible to automate more complicated labour processes in a technically simple and cost-effective manner. There is also a continually increasing flexibility in the design and application. The high speed and constant quality which numerically controlled machines have achieved with the aid of currently available programs, can no longer be matched by a conventional machine which is operated by a skilled worker.

Vocational education developments

From evaluation studies, it appears that the rate of development of vocational education has fallen behind the rate of development in industry. A major cause of this is a scarcity of financial resources (Comissie evaluatie INSP, 1988). Despite this, however, during the eighties in the various forms of Dutch vocational education, work has been proceeding on the adaption of the programmes with regard to the technological developments in the area of production automation. For the development at a regional level, the regional vocational centres for advanced technology are of particular importance for CNC-production/FPA In the second half of the eighties, six regional vocational centres for mechanics were established, while in the beginning of the nineties three more have already been started. However, over the years, ideas about the function of the vocational centres have changed. The presently existing centres are intended for businesses and schools for the purpose of additional training in advanced technologies. The current target groups are: students and teachers in higher vocational education and on-the-job trainees; graduates of mid-level vocational education; workers, the unemployed, employers, and entrepreneurs.

The place that the regional centres should have within the educational infrastructure is currently still under discussion. In contrast to a number of years ago, the majority of schools are presently satisfied with the availability and usefulness of the regional centres. From a technical point of view, the guidance is generally considered as quite professional. At present, the most problematic areas are:

- the financial aspects;
- the connection with learner's previous school curriculum;
- matters relating to lesson scheduling and teacher participation.

In the paper presentation this will be further explained.

Literature

Braverman, H. (1974). Labor and monopoly capital. New York

Comissie evaluatie INSP (1988). Evaluatie van het INSP-programma in het veld automatisering. (Evaluation of the INSP program in the field of automation.) Zoetermeer: Min. O&W

Kayzel, R. (1986). automatisering en vakmanschap. (Automation and craftsmanship). In: Tijdschrift voor arbeidsvraagstukken. 2. 1, Pag. 60-71

1064

OECD, Directorate for social affairs, manpower, and Education (1990). The Changing role of vocational and technical education and training. Paris: OECD

Streumer J.N., & Feteris A. (1991). Responsiviteit van het beroepsonderwijs ten aanzien van technische verniewingen: CNC-verspanen en flexibel productie automatisering. (Responsiveness of vocational training with regard to technological innovation: CNC-production and flexible production automation.) Enschede: Universiteit Twente/ TO/ OCTO.

RESPOSIVENESS OF VOCATIONAL EDUCATION IN DUTCH PRINTING INDUSTRY

B. Hövels, & J. van den Berg, ITS, Nijmegen, The Netherlands

Summary

The responsiveness of the vocational education in Dutch printing industry has a long tradition. This is due to the high degree of institutionalisation and regulation of labour market, vocational structure and education in this sector and due to the permanent necessity to react on the extensive and fast technological developments in printing industry.

Social partners in Dutch printing are traditionally the guiding actors supported by the GOC, the institute charged with policy preparation and execution. Recently the guidance perspective is more directed to the adaptation of education while in the past labour and vocational structure were also object of responsive efforts.

ISSUES IN DEVELOPPING CURRICULUM PROGRAMS FOR CORPORATE TRAINING AND EDUCATION

Martin Mulder, Department of Education, University of Twente, Enschede, The Netherlands

FROM RESEARCH BASED JOB PROFILES TO PROPOSALS FOR BETTER COURSES: WHAT ROLE CAN SCIENTIFIC RATIONALITY PLAY IN THAT ROUTE?

A.M. Versloot, Department of Educational Research and Department of Education, ISOR, University Utrecht, Utrecht, The Netherlands

What role can researchers play in bridging the gap between analyzing job profiles and developing relevant course content and design?

Context of the problem

The educational department of the service and research organization for the Dutch gas- and energy distribution industry, the VEG-Gasinstitute, financed a research project to describe the profiles of a number of gas-technical jobs. These profiles were meant to provide the empirical base for the reconstruction of two existing courses for employees in this industry. From March 1990 until July 1991 researchers of the Faculty of Social Sciences of the University of Utrecht executed this research and development project in cooperation with a number of staff members of the VEG-Gasinstitute.

The following factors provided the impetus for the projects' initialization:

- the rigid form of course organization;
- the abstract nature of the learning materials;
- the lack of educational professionalism of technical specialists in the role of teacher;
- the wide gap between the content of the lessons and practical problems encountered by employees in their daily work;
- new technological and organizational developments (scale enlargement) which call for more flexible employees and, as a consequence, new course contents.

This means that the gas distribution companies are faced with the well known problem of mismatch between knowledge, skills and attitudes of employees necessary at the workplace and the knowledge, skills and attitudes students are

1066

taught in the educational and vocational training institutes (Nijhof, 1986, Mulder & Thijsen, 1990). A training needs assessment must be done (Rossett, 1987). The research design can be briefly sketched as follows. After a preparatory phase of interviews with informants of 22 gasdistribution companies a questionnaire was developed for selected employees to establish the importance of task components in their jobs. The selected employees also answered questions about their job career pattern, their educational career and about the importance of a great number of knowledge, skill and attitude items for the performance in their job. At a conference we presented for a group of stakeholders of the two courses our preliminary empirical results and presented tentative conclusions and propositions for curriculum goals. And in our research report we did produce the job profiles. Can the researcher then be satisfied and trust the developers of courses to make proper use of the delivered empirical data? Trainers are all too familiar with the problem of transfer of training results. What about the problem of transfer of scientific insight in job profiles, based on empirical reseach, into better quality courses?

Problem

What role can researchers play in bridging the gap between analyzing job profiles and developing more relevant course content and design? What role is there to be played by scientific rationality? I want to stress the importance of the question of transfer of research data and conclusions to the context of solution of a practical problem. The process of choice and validation of course content is not at all ended by the presentation and acceptation of job profiles. I contend that there is multi-stage process necessary to develop or reconstruct a curriculum or a course, based on job profiles. Has the researcher a role to play in that process? Traditionally the role of the researcher is limited to the formulation of precise research questions, correct research design, professional development of research instruments, organized collection of data, efficient techniques of data-analysis and finally reporting and presenting clearly the results and conclusions. From a purely scientific perspective the formulation of recommendations would not even be part of the role of the researcher.

Empirical job analysis research pretends to make this coordination problem solvable on an rational base. Research based job analysis results in detailed job profiles, that summarize enormous quantities of data (Veen van der & Versloot, 1992). The researcher can take the attitude that his task is completed by delivery of job profiles and that the follow-up phase concerns a development process that can proceed in optimal form on the base of experience, intuition and subject matter expertise. In that case staff specialists responsible for choice of relevant contents, curriculum structure and course design must continue where the empirical researcher makes his exit.

Method to improve transfer of research results

Most of the research effort can easily go wasted if researchers don't participate in the process of innovation. Boekenoogen and Stokking (1991) report that, even when needs analysis reports are based on sound methodological and empirical methods, the results are left unused if they are perceived by development staff as irrelevant to the practical problem. The problem of transfer of research results is

caused by cultural and language difficulties between researchers and curriculum developers and by differences in responsibilities, professional ethics and interests. The transfer problem can however partly be overcome if researchers accept co-responsibility for the development process and if they are accepted by developers as worthy contributors. When the researcher, as in the Dutch Gas-project for curriculum innovation, accepts co-responsibility for the phase of curriculum development, certain actions can contribute to better transfer of research results. There are at least two important transfer points in this kind of innovation process. Firstly, the transition from job profiles to the choice and formulation of learning goals for curriculum and separate courses as components of the curriculum. Secondly, the transition from learning goals to the design of courses and choice of contents. I would like to report on the specific procedure we chose for the cooperation with staff specialists of the VEG-Gasinstitute to bridge these transition points.

- we specified the steps in the multi-stage development process on the base of the produced job profiles and additional profiles of necessary knowledge, skills and attitudes
- we accepted and welcomed a monthly get together with four curriculum developers
- we requested the willingness of the developers to evaluate provisional documents we planned to present to them
- we made clear that the empirical data base for our contribution to the developments was threefold: a. the protocols of interviews held at the gas distribution companies with a number of informants (a line manager, a member of the personnel staff and a training consultant), b. the primarily quantitative data collected by questionnaire on tasks components, job career, educational career and necessary knowledge, skills and attitudes and c. the results of the study conference held with selected stakeholders connected with the two courses that needed to be reconstructed.
- we announced a twofold strategy for our contribution to the development process. Firstly, we implemented an inductive approach based on detailed empirical results to narrow down the range of choices for important learning goals and the design of curriculum and separate courses. Global concepts on which to build a renewed curriculum don't come out of the blue sky. The inductive approach obliges the researchers to a very intensive analysis of all the available data. It seems almost unavoidable that this focused attention on details generates ideas about obsolete components in the existing curriculum and missing important components. Furthermore the intensive process of cooperation between researchers, trusting on the collected empirical data, and developers, trusting their experience and subject matter expertise, generates also unavoidably new conceptions for the design of a curriculum and separate courses. Secondly, once global conceptions were formulated for a reconstructed curriculum a deductive approach could be implemented. The global conception was specified and detailed so that separate components of the curriculum got shape and choices were made for course content, course organization, presentation of lessons and all other relevant aspects.

The multi-stage development process consisted of a very disciplined effort during each month of performing analyses, preparing discussion documents, exchanging

comments and detailed proposals and reaching consensus on important issues. We as researchers presented at the end a detailed proposal for the reconstruction of the two existing courses. We made fully use of the opportunity to clarify our purposes and intentions. But at the same time we accepted the responsibility of the development staff of the VEG-Gasinstitute to make their own final choices. In the balance we are sure that our empirical results have seriously contributed to the development process and the quality of the final results. But as researchers we had to be part of the process to guarantee that our research efforts would not go wasted. By looking at this participation in the process as an extension of traditional research activities researchers can determine on an empirical base what mechanisms of social interaction are responsible for the frequent occurrence of research reports treated as irrelevant. The implementation of research results can't be left to members of other disciplines.
Researchers cannot limit themselves to the execution of purely rational and mathematical procedures and report on the results. The rational arguments of researchers have to compete with the know how and subjective preferences of staff members. If researchers want better transfer of research results they have to take a more political, interactional perspective and their methods and activities accordingly.

References

Boekenoogen, M.S., & Stokking, K.M. (1991). Methoden van behoeftenonderzoek. Utrecht: ISOR.

Mulder, M., & Thijsen, A. (1990). Decision making in curriculum conferences: a study of convergence of opinion. Journal of Curriculum Studies, 22, 4, 343-360.

Nijhof, W.J. (1986). Beroepsprofielen en curriculumprofielen; tussen aanpassing en innovatie. In: Nijhof, W.J. (red.), Van beroepsprofielen naar curriculumprofielen: een symposium over methoden en technieken en beschouwingen over de relatie tussen beroepskwalificaties en eindtermen. Enschede: Faculteit Toegepaste Onderwijskunde.

Rossett, A. (1987). Training Needs Assessment. Englewood Cliffs (NJ): Educational Technology Publications.

Veen, J.M. van der, & Versloot, A.M. (1992). A Research-based model for job-analysis. A case from Dutch Gas- and energy distribution industry. Journal of European Industrial Training (in draft).

TOWARDS DESIGN STANDARDS IN CORPORATE EDUCATION

Joseph Kessels, Foundation for Corporate Education, Terschuur, The Netherlands

Introduction

This paper contains the first report of a multiple case study on 8 successful and 8 non successful training programs. Major research questions are:

- Does a structured program design enhance the quality of the program?
- What factors influence a successful implementation?

Conceptual Framework

The study is based on the construction of a curriculum typology for corporate education (see chart 1). The basic elements of this typology are derived from Goodlad, Klein & Tye (1979). However different labels have been used and much emphasis is put on the consistency between the subsequent curricula.

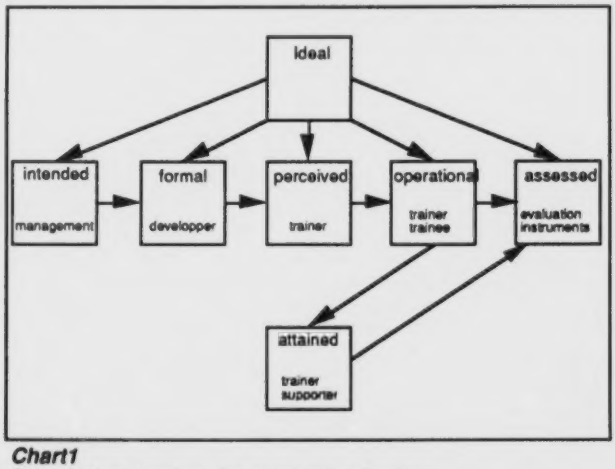


Chart1
A curriculum typology for corporate education.

Ideal curriculum

The ideal curriculum is the most adequate set of learning situations that contributes to solving a given performance problem in the company.

Intended curriculum

The intended curriculum contains the assignment given by (top-)management to the training department. Management should also state what it will do to create

favourable conditions in the workplace, for the acquired skills to be transferred into the desired performance.

Formal curriculum
The formal curriculum contains all documents, training materials, written instructions, schedules, excercises, software etc. of the training program which have been produced by the curriculum developer(s).

Perceived curriculum
The perceived curriculum reflects the trainers perception of the curriculum.

Operational Curriculum
The operational curriculum is the curriculum as it presents itself in the real learning situation. It consists of the factual learning situations that are created, the interaction between trainer and trainee, the trainee working with the curriculum materials, and the learning processes that take place.

Assessed curriculum
The assessed curriculum contains all the formal outcomes of a training program that are recorded by means of evaluation instruments. When the evaluation instruments are well designed, they show: the reactions of the trainees on the training program, the learning results, the changes in performance in the workplace, and the impact of the training program on the organization (Kirkpatrick, 1975),(Hamblin, 1974).

Attained curriculum
The attained curriculum consists of all effects that are caused by the training program. The attained curriculum should incorporate the solution to the problem that was previously stated in the intended curriculum. Even if the training program brings about the desired changes in skills, compentencies and abilities of the trainee, this doesn't mean automatically that these changes have an impact on the organization, (Robinson & Robinson, 1989).

Towards instructional design standards in corporate education
The curriculum typology as described above allows us to operationalize the concept of quality of corporate education. As the ideal curriculum states the most adequate contribution of training to solving a problem in the organization and as the attained curriculum states the factual training results and their impact on the organization, the quality of corporate education is of high level when the attained curriculum reflects the ideal curriculum. This means that the quality of corporate education can be expressed as the consistency and discrepancy between ideal curriculum and attained curriculum. The basic elements of the ideal and attained curriculum, as well as the relationships that influence consistency and discrepancy are shown in chart 2.

Research Design

The study was carried out as a multiple case study, with multiple units of analysis, focussing on theoretical replication, discribed by Yin as a Type 4 - study, (Yin, 1989). 8 trained data collectors conducted document search and had interviews with training managers, line managers, supervisors, developers, trainers and trainees. Each case study report was coded seperatedly by two different data analists. The codingschemes were derived from the research questions. Within-site analysis and cross-site analysis were carried out by means of display techniques as discribed by Miles and Huberman (1984).

The following aspects were subject of analysis:

- The consistency between actors' perceptions of the initial problem to be solved.
- Actors' involvement during development and implementation.
- Program design.
- Consistency between assignment, objectives, evaluation, learning situation and training materials.
- Cost-Benefit Analysis
- The consistency between actors' perceptions of results and effects.

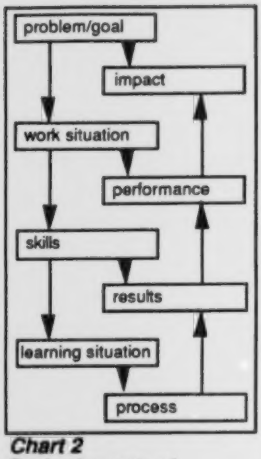


Chart 2
Basic elements of Corporate Education

Conclusions

- In the 16 cases examened a well structured program design is rarely found. The consistency of curriculum elements and subsequently the quality of the program does not depend exclusively on a structured program design.
- Successful implementation of a training program largely depends on the following compensating factors:
  Involvement of line managers and supervisors;
  Close similarity between the learning situation and the workplace;
  Trainer has gained practical experience in the subject matter field.
- Needs assessment is rarely carried out, and poorly documented.
  Evaluation criteria are nor documented nor stated.
  Evaluation instruments are resticted to the lowest levels of evaluation.
  Subsequently cost analyses are rare. Analyses of benefits have not been found.

The results of these case studies have been transformed into improved design instructions. These instructions focus strongly on a structered design, with emphasis on the above mentioned compensating factors (see #2) and on implementation activities *during* the design process (Plomp, 1982). These design instructions will be validated in an additional study of 30 training development projects carried out in 30 organizations. Results will be analyzed in the second half of 1992, which might lead to improved instructional design standards for corporate education .

1072

References

Goodlad, J.L., Klein, M.F., & Tye, K.A. (1979). The domains of curriculum and their study. In: J.L. Goodlad et al., Curriculum inquiry: The study of curriculum practice (p. 43-76) New York: McGraw-Hill.

Hamblin, A.C. (1974). Evaluation and control of training. London: McGraw-Hill, 208 p.

Kirkpatrick, D.L., (1975) Evaluating training Programs. A collection of articles from the Journal of the Americam Society for Training and Development. Madison: ASTD, 313 p.

Miles, M.B., & A.M. Huberman, (1984) Qualitative Data Analysis. London: Sage. 264 p.

Plomp, Tj., (1982). Onderwijskundige technologie: enige verkenningen. Twente: TO. 60 p.

Robinson, D.G., & J. Robinson, (1989). Training for Impact. San Francisco: Jossey-Bass. 308 p.

Yin, R.K., (1989). Case study research. Design and Methods. London: Sage. 166p.

1073

NEW TECHNOLOGIES, SHARED CONTROL AND (UNION-) EDUCATION

C.J. Feijen, & B. v. Onna, Katholieke Universiteit Nijmegen, Nijmegen, The Netherlands

In order to cope adequately the economic increasingly difficult situation, companies all over the world try to improve their competitiveness by introducing several types of new technologies. Des-pite of the possibilities of these technologies to increase company-effectiveness, the introduction of computer-based systems often leads to social consequences which are undesirable from the perspective of employees. This problem has evoked initiatives from the unions and workers' re-presentatives in companies, to take care of the interests of people on the 'shop-floor', who ultima-tely have to work with the systems. To be able to stress the needs and interests of the employees in a more or less fundamental way, many workers' representatives try to get involved in a techno-logical project as early as possible. Practice has shown that such an approach is very difficult. Strategies to take proper action in the separate phases of a technological project are scarce. Lots of workers' representatives expect or need educational support to increase their competencies in order to serve the interests of the employees as good as possible if technological changes are in-tended. But in the educational setting there is not enough luggage to give the support that wor-kers' representatives need to operate effectively. The shortcomings within the educational scene, already known for several years, were the point of departure for a project called 'Action structures of workers' representatives and educational supply'. The target of that project was the formulation of educational guidelines as a result of an intensive empirical research-period.

The research was built on two fundaments: the psychological action theory of Hacker and Volpert, that does not only stress the visible aspects, but also the - planned - intellectual regulation of ac-tions, and a division of the process of shaping and introducing new technologies into several mo-re or less well defined phases.

The first part of the research was aimed to receive notions about desirable action structures - as possible educational goals - for workers' representatives that can be activated in the several pha-ses of the total technological project. Twenty experts [from Germany and the Netherlands] were submitted to a semi-structured interview that lasted from 2 to 4 hours each.

The second part of the research consisted of extended - structured - case studies in 5 companies. In each company at least 4 workers' representatives were interviewed, 1 or 2 to get a good picture of the context in which they had to operate and 3 to examine exactly the actions taken by the wor-kers' representatives collectively and individually to make sure that certain interests of the working people involved in a concrete project were not ingnored. In the 5 companies also employees, the users of the technical systems were approached. They have filled in a questionnaire about user-participation and the interests of employees. In fact the case studies were meant to investigate the actual action-structures of workers' representatives.

1074

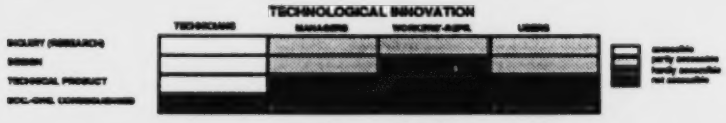
The third part of the research existed of an elaborated questionnaire that was sent to 275 compa-nies, to check the representativeness of the action-structures found in the 5 cases, and to test the value of a few central research theses.
The research-part in which the desirable action structures were focussed, has resulted in six -part-ly combinable- structural types. The first structural type is the input-enrichment, implicating that workers' representatives develop norms from a social-organizational point of view (if possible al-ready at the beginning of a technological project) that should be taken into account by the develo-pers within the project, who usually only are concerned about the technical side of the coin.
The second structural type is the transformation-control. That means that workers' representatives try to break down the too narrow technical orientation, not by formulating social-organizational norms, but by critically studying the functional specifications in order to steer (modify) them in a socially more convenient direction. Within a transformation-control workers' representatives try to execute the translation step from the social to the technical sphere by themselves, in stead of lea-ving this task to the technicians.
The third structural type is the shaping of the technical core-concept. That implicates that starting from *concrete* technical options and *concrete* organizational alternatives, a future work situation is shaped, in which the interests of all parties involved are optimally integrated.
The fourth structural type is the shaping of the technical details, implicating that workers' repre-sentatives (when the more fundamental decisions already have been made) try to create a situa-tion in which a user-design can be developed that properly links up to the existing working-me-thods and meets the whishes of the employees who ultimately have to work with the technic(al system).
The fifth structural type is the shaping of the organization of work (task divisions). That means that workers' representatives mainly concentrate on alternative organizational options that, given a certain kind of technic, still can be realized.
The sixth and last structural type is the output-softening, implicating that workers' representatives (al-most) at the end of the developmental process, try to alleviate some negative consequences of the technical-organizational choices already made.
What workers' representatives should or could do to implement the various structural types, and in what way education can be supportive, is described in detail in a (my) thesis, that contains for each structural type action plans, information about necessary - mostly not available - competen-cies and information about structure-specific advantages and disadvantages. The structural types one, two, four, five and six can be performed within the most dominant system-introduction-metho-dologies used in practice. The structural type three demands a fundamental other way of introdu-cing new technologies.
The research part concerning the actual action structures has made clear that workers' represen-tatives normally take care of the interests of employees when new technologies are introduced, by means of the structural type output-softening and especially the structural type input-enrichment, that both is implemented apart and in combination with shaping and output-softening activities. The structural types two, three, four and five have not been found among the workers' representa-tives investigated. Characteristic of these types is a stronger contential

commitment that, as the re-search concerning the actual action structures also showed, most likely leads to better results than a more conditional approach. Workers' representatives can improve their acting in relation to new technologies by the optima-lization of the structural type they implemented untill now, by the replacement of this structural type by another - more promising - structural type, or by the combination of their current structural type with one or more other structural type(s).

In principle workers' representatives can, roughly speaking, take the next - separate- courses. They can go on formulating conditions (norms) that in the second instance have to be made ade-quately concrete by the technicians. They can try to get involved in the formulation of the functio-nal specifications, a practice that in many cases is only possible when workers' representatives can activate technical experts for their own goals. They can wait untill the innovation has got a more concrete shape, to intervene in the creation of the technical details and/or the necessary c.q. whishfull organizational adjustments. They can let the technical-organizational innovation pass, and try to soften the negative consequences of this innovation as good as possible. Or they can concentrate their energy on changing the top-down system-introduction-methodology into a bot-tom-up approach.

The last option - the third structural type - seems to be the most suitable. That option implicates that employees, technicians, management representatives and workers' representatives develop concrete future work-places, with the technological possibilities integrated satisfactorily, not only to the benefit of the company but also to the benefit of the employees involved.

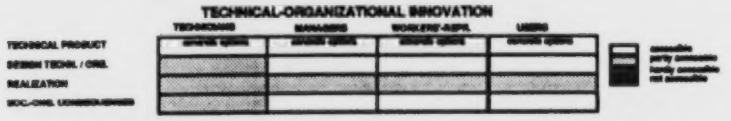
Technological innovations nowadays mostly are achieved as follows. At first an inquiry is execu-ted. On the basis of the information from this inquiry, mainly management representatives (among which eventually also some users) and technicians make an abstract design. That design then is being transformed into a technical product by the technicians. And (almost) all at the end of the process it becomes really clear what the consequences of all efforts are. The problems with such an approach can be demonstrated by the following outline.



For the technicians the course of events in a project is the most accessible. They mainly lack an adequate insight into the social-organizational consequences of their activities. Workers' repre-sentatives are more than the other parties in the dark. They probably can gain a certain insight into the inquiry, but the following steps mostly are unclear for them, so it is difficult to reach a sub-stantial amount of influence. Managers and users are in a position in between. For them the first two steps are partly accessible, but because afterwards almost only the technicians decide what is being done, they too are not able to get a good picture of the social-organizational consequen-ces of the renewals.

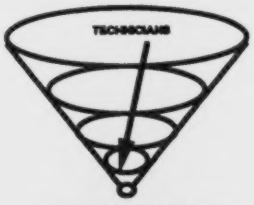
1076

Strictly speaking a technological innovation should give way to a technical-organizational innova-tion, starting with concrete options in stead of abstract intentions. Then the next picture (outline) would become dominant.



A technical-organizational innovation usually is much more transparant. The reason for this is that the technical-organizational design is made on the basis of concrete technological options, that in a more or less direct way refer to organizational models and their possible alternatives. The effect of a more concrete approach is a (relatively) fast knowledge of the several (dis-)advantages that go along with a certain kind of renewal, for the company as well as for the employees.
A 'technological innovation' as described above, has a certain 'bow-net'-effect. Decision proces-ses occur within ratio X (the digital ratio) that from ratio Y (the analogous ratio) hardly can be in-fluenced. In short, non-technicians are scarcely able to take care of an optimal unity of the social desirabilities and the technical possibilities, if they are confronted with a top-down system-intro-duction-methodology that is characteristic for a technological innovation. That unity can only be shaped when the digital bow-net will be replaced by an analogous bow-net, see below.

DIGITAL BOW-NET



ANALOGOUS BOW-NET

Not the technical formal (digital) logic should subordinate the much richer practical (analogous) logic, but the practical logic should predominate. Only in this way the technic really can prove its mean-character.

KNOWLEDGE UTILIZATION WITHIN VOCATIONAL EDUCATION ENVIRONMENTS

W.J. Nijhof, Department of Educational Technology, University of Twente, Enschede, The Netherlands

Abstract

The demographic trend and the impact of new technologies change jobs and organisational structures in enterprises and administration. As the new tasks are more complex, an adequate preparation for the worksite is needed. But modes of traditional instruction do not seem to fit into the gap between general as well as vocational education and qualification required by the labour market: School follows its usual pattern; work processes in production and administration have become and are still becoming more complex and risky that introductory courses must be taken out of the productive processes. New forms of knowledge acquisition and utilization are needed. One solution of this problem may be to develop simulations, especially in the form of complex teaching-learning environments or new forms to train knowledge utilization. The contributions to this symposium, therefore, demonstrate new ways of knowledge acquisition and utilization in natural setting in the field of vocational education. By using hypotheses won in the fields of cognitive psychology, theory of action, didactics and corresponding sciences (here: business administration) and also by using computer facilities, theoretical and practical experiences will be reported and evaluated.

Objectives of research

- The demographic trend in mostly all European industrialized countries leads to a shortage of workers and clerks. There is or will be a gap between the needs of the worksite and the number of trained people.
- The impact of the new communication and information technologies changes jobs and organisation - mainly in direction to more complex task structures.
- These task structures on the worksite - in short - are distinguished by - a large number of variables with different degrees of transparency - which form a network, and - vary over the time (also non-linear), and - are polytelic (the problem of main- and side-effects of decisions).
- By school education, but especially by vocational education, students should learn to treat such problems, to acquire and utilize the correspondend knowledge, and to judge the quality of problem-solving processes.

1078

- At the moment, school is not prepared to adequately handle these tasks. This fact makes it necessary to develop teaching-learning environments which fit into the gap between general as well as vocational education and qualification required by the *labour market.*
- But also enterprises are confronted with difficulties to adequately train their workers and clerks. Since processes in production and administration have become and are still becoming more complex and risky, introductory courses as well as courses of recurrent education must be taken out of the productive processes.
- Therefore, the teaching-learning processes are assigned to schools or simulation processes which serve as models for the real situations.

Theoretical framework

- The symposium will be dealing with the wide range of possible preparation for the worksite:
  Achtenhagen gives an overview and reports on the wide range of teaching-learning arrangements which are used to prepare for fulfilling the complex tasks in the field of commerce. Such arrangements which can be used per se or in combination are simulation games, case studies, working-analogue learning tasks, simulated offices, special training procedures and so on. They are discussed as part of the *formal education* within vocational schools.
  Mandle shows - using the simulation game presented by Achtenhagen - how such tools can be used to build up mental models. He concentrates on (a) utilization of declarative knowledge (b) utilization of external "tools", (c) problems of misleading concepts, (d) interaction between pairs.
  Neber demonstrates the relationship between content specific knowledge and problem solving. He shows effects of the utilization of knowledge in the field of (technical) vocational education.
  Achtenhagen demonstrates possibilities of a schema-based modelling of complex economic situations. This approach - a computerbased simulation game enables to thoroughly describe content and objectives of teaching-learning environments in relation to the curriculum of a *commercial school* as part of the *formal vocational education.*
- The symposium, therefore, demonstrates effects of complex teaching-learning environments within different institutions. In combination with the systematic discussion of teaching-learning theories and problem-solving strategies, advantages and disadvantages of different models of "reality" for developing worksite-related qualifications can be shown.

Methods and techniques

In the papers, the following points are discussed:

- Description of different types of simulation;
- Development of measures of the distances between "reality" and the environment used;
- Schema-based modelling of environment;
- Relationship between types of simulation and authentic activities;
- Description of teaching-learning and problem-solving processes within different environments;

- Measurement and diagnosis of learning outcomes (including skill acquisition).

Research tools are
- Long- and short-term observation and description of teaching-learning processes and environments;
- Protocols of group discussion during task-fulfilling;
- Questionnaires;
- Tests;
- Computer-based methods of learning diagnosis.

Data analysis
The data are analysed and discussed according to the different research methods and techniques. Qualitative and quantitative methods are used to describe the strength of relationship between different types of teaching-learning environments and learning outcomes.

Results
The introduction of complex teaching-learning environments and training method by modern enterprises and schools to adequately prepare students and apprentices for the worksite demonstrates that entrepreneurs and trainers as well as headmasters and teachers very strongly believe in these methods: They offer a chance for effectively teaching and learning. By the research discussed at this symposium, that impression shall be evaluated and fostered.

Educational Importance
Confronted with the changes at the worksite, school as well as vocational education have to react. The use of complex teaching-learning environments seems to be a very effective alternative to "normal" educational procedures. Therefore, the results of this symposium will offer interesting and important hints for actual practice and future research.

COMPLEX TEACHING-LEARNING ENVIRONMENTS- THEORETICAL ASSUMPTIONS AND PRACTICAL DEVELOPMENTS

Frank Achtenhagen, der Georg-August-Universität, Göttingen, Germany

Objectives
- The paper especially deals with vocational education on the field of commerce.
- The impact of the new communication and information technologies causes substantial changes of organisation and tasks at the worksite.
- These changes lead to new forms of task enrichment in the fields of economy and commerce.

1030

- By that, problems to be solved also change and become more complex. This complexity - in short - is distinguished by:
  a large number of variables with different degrees of transparency - which form a network, and
  vary over the time (also non-linear), and
  are polytelic (the problem of main- and side-effects).
- In Commercial Education which directly prepare for working in offices and administration, students should learn to treat such problems and to judge the quality of problem-solving processes.
- Objective criteria are - according to economic theory - e.g. figures of productivity, economic efficiency, profitability, liquidity, which can be counted by given equations.
- These criteria are fixed as goals of the curriculum in the first year of German Commercial Schools (tenth year of school: 9 year of general school and 1 year of commercial school). They also are accepted by the employers' association.
- According to the changed needs at the worksite, we developed a program to teach these goals and content by simulation processes (we are sponsored by the Federal Ministry of Education and Science and the State Ministries of Education resp. Science).

Theoretical framework
We combined approaches of cognitive psychology, theory of action, and didactics on one side and the system-oriented theory of business administration (developed in St. Gall, Switzerland) on the other. Doing so, we constructed a series of complex teaching-learning environments: simulation games, case studies, working-analogue learning tasks, simulated offices, per se and in combination - with use of computers.

Methods and techniques
We analysed the teaching-learning processes for several months - by videotaping and tape-recording the discussions and interactions in the classroom and in groups. We also administered questionnaires and criteria-oriented tests. All results are taken to revise the curriculum and the teaching-learning arrangements. We now control the instructional processes in the fifth year. According to this procedures of development and evaluation, our curricula will be officially implemented and disseminated by the ministry of education.

Data analysis
Selected data will be presented which show that it is not enough only to develop new complex teaching-learning environments. It is also necessary to develop strategies to adequately handle the processes released by these environments. The data are concentrated on strong side-effects of the teaching-learning arrangements which injure the intended maineffects. The data are won by protocol analysis, test analysis etc.

Results
The results demonstrate that by complex teaching-learning arrangements the understanding of complex task structures can be fostered. The new environments seem to lead to better learning outcomes than traditional instruction does.

Educational importance
Ways are shown by which the youth in industrialized countries effectively can be prepared for there worksite. As we try to also stabilize the development of students' personality - by the active, self-governed authentic learning processes -, the assimilation to needs of the labour market is balanced by accommodation processes.

SCHEMA-BASED MODELING OF COMPLEX ECONOMIC SITUATIONS

Frank Achtenhagen, der Georg-August-Universität, Göttingen, Germany

Objectives
The impact of new information and communication technologies changes the structure of economic and commercial tasks in offices and administration. This process requires a change of traditional patterns of commercial education in schools and enterprises. The most promising way is to introduce simulations in the form of complex teaching-learning environments. To adequately handle the simulation processes, it is necessary to develop methods for measuring the distances between "reality", and the model. One way may be to describe the processes of reduction and complexion and to make them transparent.

Theoretical framework
The description of modeling processes is done by using schema-based knowledge representations. The development of this system uses a combination of hypotheses within cognitive psychology (the schema-based approach of conceptual dependency) on one hand and within a system-oriented approach of business administration on the other. By steps of modification, dimensions of semantic primitives could be developed. These primitives form the framework for a description of actions.

Methods and techniques
A specific teaching-learning environment: a computer-based simulation game, is described by this model. The units of action and reflection can be combined and brought into a hierarchy of cognition. By this, the planning of instructional units is strengthened and can be evaluated.

Data analysis
The different levels of reduction resp. complexion processes which form the basis of simulation of "reality" are fixed. The model allows to mark the differences. As the model also can be used to describe learning outcomes, the structure of the

1082

environment, of the teaching-learning processes and the test results can be related to each other.

Results
The results demonstrate - for different learning groups - that this procedure can fit teaching-learning environments into the prior knowledge and the learning capacity of students and apprentices.

Educational importance
The modeling procedure seems to be applicable not only to simulation games but also to other examples of complex teaching-learning environments. The results are won within a compulsory part of German commercial education.

KNOWLEDGE UTILIZATION: COMPLEX PROBLEM SOLVING IN COMMERCIAL EDUCATION

Heinz Mandl, Harls Gruber, & Alexander Renkl, der Ludwig-Maximilians-Universität Munich, Munich, Germany

Objectives
- In school, very often, the goals and content are taught in a linearized form.
- There are four connections between the different parts of declarative knowledge.
- On the other hand, the new impact on the labour market urges a knowledge which is complex and delinearized.
- By using a computer-based simulation game effects of complex knowledge acquisition and utilization are tested.

Theoretical framework
Research goal is a detailed analysis of cognitive and social processe.s during the utilization of knowledge for complex problem-solving. Theoretical background is the concept of "mental models".

Methods and techniques
We investigated students of a German commercial school. By a test of prior economic knowledge, pair of students with differing outcome measure were set. These pairs, then, had to run a computer-based simulation game of jeans-fabrication over six periods. The students had to think loudly and had to express arguments for their decisions.

Data analysis
The test data, the protocol data, and the outcomes of the simulation game were compared.

Results
Conditions and reasons for success and failure were identified.

Educational Importance
Lacks of knowledge utilization could be identified. The outcomes show ways how to handle these problems. Therefore, important hints for the development of processes of knowledge utilization in the field of vocational education were marked.

KNOWLEDGE UTILIZATION: PROMOTION BY A TRAINING IN SITUATED ACCESS

Heinz Neber, Universität Essen, Essen, Germany

Objectives
- The utilization of knowledge is a central point of educational psychology.
- This problem is characterized as "inert" knowledge; it is central for research on transfer.
- The utilization of knowledge in rich environments, e.g. in the worksite, is therefore important for teaching-learning processes in the field of vocational education.

Theoretical framework
It is shown that the way of knowledge utilization is changing: explicit, automatic (with trials), and automatical use of knowledge. These steps mark the development from novice to expert behaviour. The concept is used for research on diagnostic behaviour.

Methods and techniques
Students were investigated on their behaviour of formulating hypotheses and planning. They received external help by lists of possible faults of motors. 60 apprentices (car mechanics) were tested: prior knowledge, diagnosis behaviour (with different conditions). Then, they received different training on knowledge utilization.

Data analysis
The complex design was analysed.

Results
Different ways of knowledge utilization - under the different training conditions - could be identified.

Educational Importance
Different effects of training are shown. These effects can be used to optimize teaching-learning processes in (technical) vocational education.

1084

EDUCATION AND TRAINING FOR LOWER EDUCATED ADULTS

T.F. Brandsma, Centre for Applied Research in Education (OCTO), University of Twente, Enschede, The Netherlands

Introduction

The attention of educational policy and research for continuing education and training for adults is not new in itself, but has changed its character during the last years. More and more the economic necessity of continuing (vocational) education and training is emphasized. Recent initiatives in EC-context, like the FORCE program (established on behalf of the promotion of continuing vocational training in companies throughout the European community), underlines this.

However, many studies in the field of (non)participation in adult education prove that, those that benefit from continuing education and traning are the better educated, with a good job position and a high(er) social-economic status. It appears that adult education does not, as was once hoped, reduce educational inequality. Various theoretical notions have been elaborated to explain the phenomenon of the limited participation of lower educated adults. On the one hand the relationship of initial educational attainment, labour market position and personal situation with (non)participation in adult education is stressed. On the other hand explanations are sought in the (limited) possibilities of access to continuing education and training, the characteristics of educational programs thath are provided and the discrepancies between these programs and the lower educated adults' needs.

This symposium concerns the (continuing) education and traning for lower educated adults This theme is approached from two viewpoints:

- the (lower educated) adults' needs for and preferences with regard to adult education and training as well as their participation in specific activities;
- the deterrents and barriers which inhibit these adults' actual participation and the (possible) solutions for reduction of these deterrents and barriers.

The contribution of Van der Kamp and Janssen considers the content of adult basic education (especially basic skills that should be acquired) and compares the judgements of participants in adult basic education with those of experts on this topic. It appears that there is a considerable gap between the basic skills participants stress as important and those that experts emphasize. Doets' contribuiton goes in on the determination of (the nature and size of) adults' needs for adult basic education and the consequences of these needs for educational investment.

1085

The contribution of Moerkamp forms a bridge between the first and the second viewpoint; she presents the results of a study concerning the number of participants in adult basic education that move on to further education and the problems these participants meet once they have entered these courses. Onstenk presents in his contribution the results of a study in which the actual participation of older workers with low educational qualifications was analyzed as well as the (succesful) characteristics of specific (in-company) training projects for this group of employees. The contribution of Ruijs presents recent data on the participation of specific target groups in adult basic education and analyses the influence of the international literacy compaign and the policy accent on the preparatory function of adult basic education for continuing vocational education, on the participation in various programmes. The last contribution (Brandsma) elaborates on four explanations for the (non)participation of lower educated adults in continuing education and training and sketches, based on these explanations as well as experiences in other European countries, the consequences for educational policy.

BASIC SKILLS AND AUDLT BASIC EDUCATION

M. van der Kamp, University of Groningen, Groningen, & T. Janssen, Centre for Educational Research (SCO), University of Amsterdam, The Netherlands

Background of the study

In common with most other European countries and the United States there is a provision for adult basic education in The Netherlands. Adult Basic Education (ABE) has been started in 1987 and continues former activities for minority groups, literacy programmes and general programmes on a basic level. The participants of adult basic education are people who have hardly received any education such as certain groups of immigrants, most of them are Turkish or Maroccan. Fifty percent is younger than 35 years old. About two out of three participants had no paid job. The content of adult education is restricted to those basic skills which are necessary "to cope with daily life", according to the ABE-law. So, the participants learn to read and write, get some arithmetic and social knowledge and skills. The concept of "able to cope with daily life", however, is rather vague and needs to be reflected and investigated empirically for curriculum development and evaluation.

Aim of the study

The aim of the study was to get insight into the opinions of participants in ABE as well as experts on the relevance of basic skills in the field of reading and writing, arithmetic and social skills for basic adult education.

Theoretical framework

The issue of `basic skills' (or in literature often `literacy') can be approached from various viewpoints:

1086

- the personal view of the participant in ABE;
- various domains of functioning in society such as citizenship, workplace, privacy or study;
- the body of knowledge within different subject areas.

The approach followed in this study was a combination of the three different viewpoints. So, a matrix was constructed with 69 items, consisting of daily-life tasks concerning different subjetcs as well as domains of functioning. Items were for example: `making notes during a lesson' (writing/study); `using a pocket calculator at work' (arithmetic/work) and `communicating with the teacher of one's children' (social skills/privacy).

Design of the study

In the first part of the study 40 participants in ABE were interviewed with the help of the list of 69 items. The participants were asked to judge the items following the criteria: frequency, command of the skill and relevance to learn in ABE. They could also complete the list with other skills they wish to learn. In the second part of the study 6 panels (in totally 160 persons) were asked to judge the same list of items on the same criteria. The various panels were recruited from: (a) institutions for ABE, (b) educational supporting provisions, (c) educational polic makers, (d) secondary (adult) education, (e) employers and vocational counsellors and (f) social services.

The psychometric quality of the instrument was satisfactory (alpha's > .80). An analysis was made on the discrepancies and consensus between groups.

Results

The different groups of experts revealed a remarkable consensus. Great discrepancies, however, were found between the participants on the one hand and the experts on the other. So, the experts accentuated the relevance of basic skills in the domain of work, while many participants thought these less important (perhaps due to the fact that a lot of them had no -orientation on- work). Further the experts emphasized a broader spectrum of coherent basic skills, while the participants stressed the need of more specific skills in the field of writing and reading. Knowledge of political programmes, elections, participation at work was not a priority for the participants, while the experts think this is very important to offer in ABE.

Discussion

The results of this study could contribute to the debate on aims and content of adult basic education in The Netherlands. On the one hand ABE has clear social-cultural aims, on the other hand ABE is seen more and more as an instrument to get the elementary skills necessary to follow vocational training. There seems to be some tension between advocates and opponents of these different functions. It is also good to realize that there could be graet discrepancies between the perspective of experts and the participants themselves.

NEEDS ASSESMENT ADULT BASIC EDUCATION

C. Doets, National Institute for Research and Development in Adult Education (SVE), The Netherlands

In commission of the Ministry of Education and Sciences the SVE conducts a study, whit the main goal to gain insight in the need fo adult basic education. The study started in november 1991 and will be finished towards the summer of 1992.

The study focusses on the following research questions:

- How can the nature and size of the Dutch adults' need for adult basic education be determined;
- How can, on basis of this need for adult basic education, the educational investment be estimated, that is necessary to meet these needs;
- Which (background)data can be used as indicators for this purpose.

In agreement with the Ministery, hte choice was made to focus the research on the present population of participants in adult basic education. The following working definition of the concept `educational need' is used in the study; "the discrepancy between the skills that some one at least needs to be able to function in the present Dutch society and the actual skills that the same person masters".
The research concentrates on those skills, which are (in principal) taught in adult basic education: the literacy, numeracy and social skills that a person needs to be at least able to function in private, societal and labour situations.
For the determination of the necessary skills use is made of recent research on basic skills in adult basic education and of recently by the SVE formulated learning objectives (and the with these distinguished levels) for adult basic education.
For the research a representative sample was taken from all the participants that started adult basic education in the second half of 1991. About 3500 adult students, spread over nearly 40 institutions, participate in the research.
The respondents are asked to judge their own functional skills in the field of Dutch language, numeracy, social knowledge and skills and (elementary) English. In addition to this judgements from teachers are collected and some `objective' tests will be taken. Finally the respondents (students) are asked whether they want to learn (better) the distinguished skills.
Based on the observed discrepancies between actual and desired skills, an estimation will be made of the educational investment -in the sense of instruction and learning time- that is at least required for attainment of the desired level.
During 1993 the respondents will be approached once again, to be able to determine the outcomes of adult basic education. In the paper the first results of the needs assesment will be presented and the practicability and quality of the constructed instruments will be discussed.

1088

FROM ADULT BASIC EDUCATION TO FURTHER EDUCATION

T. Moerkamp, Center for Educational Research (SCO), University of Amsterdam, Amsterdam, The Netherlands

Introduction

In the paper results are given from a research project, commissioned by SVO and the Dutch Department of Education. Aim of the project was to get insight in the number of participants of Adult Basic Education (ABE) that move on to further education (general and vocational) and in the problems they meet once they have entered these courses.

In the research project three kinds of methods were used: case studies, a questionnaire survey and interviews by telephone. In five regions experts involved in education and training of unemployed and poorly educated adults were interviewed. Then a sample of fifty institutes for adult basic education and about seventy institutes for further education were questioned about their activities for 'transfer-students', about problems they meet and measures they think are necessary to solve these problems. Finally about 650 former participants of basic adult education were interviewed by phone (Moerkamp a.o, 1991).

Context of the study

Institutes for ABE arrange educational activities for poorly educated adults. The main aim of these activities to teach those skills that are necessary to cope with daily life situations (private life, citizenship, working life). In specific terms: literacy skills, arithmetic and social skills are trained. Preparing students for further education is part of the main aim. A few years ago it was not an important one. But recently the interest in preparing ABE students for further education has increased. Especially the Government has 'discovered' ABE as an important instrument for economic and labor market policy. According to the employment officers, many of the present unemployed, lack certain basic skills. To make them employable and 'trainable' (vocational education, training on the job), an ABE course can e important. Since 1989 the Government finances courses for long-term unemployed organized by ABE (Programs Additional Schooling, PAS). The aim of these PAS-courses is to prepare participants for vocational courses and work. So, beside the 'traditional' preparation activities, many ABE institutes organize transition courses for unemployed.

The research project not only focused on the specific transition-courses. Also transfer from other ABE-courses was studied. The number of students moving on from ABE to further education is small: less than 5% of all ABE-school leavers is actually moving on. But we should forget this percentage quickly. Most students in ABE never had plans or ambitions to move on. They participate in these courses for socio-cultural reasons, for personal development, for social contacts. For other students moving on is not realistic, because they participate in very low level literacy courses. The study was limited to students in those courses from which transfer to further education could reasonably be expected: transition courses, Dutch language courses for immigrants, arithmetic courses, etc.

A sample of these students was questioned. It appears that 38% of them had actually moved on to further education, 62% did not (at the moment of questioning). The transfer students differ from students who did not transfer, in sex, ethnic and age characteristics. Comparatively speaking, more young, male and immigrant students actually moved on to further education.
The school leavers questioned participated in different courses. The transfer percentage of those who participated in transition courses, in dutch language courses for immigrants and in arithmetic courses is above the mean percentage: about 50% of those students actually moved on.
Of the ABE leavers who moved on 46% participated in adult vocational education, 21% in adult general education and 21% in training on the job and private schooling. Particularly the older and female students moved on to adult general education. The first main reason (about 20%) not to participate in further adult education after leaving ABE is: not feeling the need. About 20% did not move on, because they found a job. Two other important reasons were: to busy with domestic work and health problems. The last two reasons are particularly given by respectively female students and older students. The most important reason to participate in further education is to improve chances to find a job (about 50%). Especially the young, immigrant and male ABE leavers are highly motivated to find a job. Another important reason to move on is to broaden one's general knowledge. Particularly for the older and female students this seems to be the main reason.
Participants of ABE who entered courses in further education have to deal with several problems. Around 60% of the students mention specific problems like: understanding, reading and writing the Dutch language, the high level of certain aspects of the course, the strict rules, lack of self discipline, too much homework. Participants in technical courses have problems with technical terminology. Particularly female students with young children have difficulties combining domestic work and schoolwork.
What do institutes for ABE and institutes for further education think about the participation of ABE students in further education? In the opinion of ABE institutes transfer of their students to further education is important. But they underline the other important aims of ABE: socio-cultural functions, eliminating illiteracy, helping students with daily life problems, preparing students for volunteer work. A majority of the ABE students is not interested in finding a job or entering further education. Therefore ABE should not be limited to economic of labor market functions. ABE institutes and institutes for further education, agree about the necessity of narrowing the gap between ABE and adult vocational and general education. They want to gear to each others targets: what do ABE students already know and what do they have to know to be successful in further education? In the opinion of institutes for further education the main function of adult basic education is teaching language skills and social skills. In the view of some of the institutes for further adult education, students who visited ABE have more social competence and are more trainable. ABE-institutes would like to organize transition courses in cooperation with institutes for further education.

1050

Conclusions

Adult Basic Education is an important provision for (very) low educated adults. About 80% of the participants received elementary education or less (CBS 1991). So it seems possible to motivate these groups to participate in adult education. But only for a little part of the participants ABE can be seen as a first step in a (new) educational career. A majority of the ABE participants don't feel the need for further education. But a clear distinction between those who want to participate in further education and those who don't, can not be made. Some ABE participants develop a need for further schooling during the ABE courses. They get more self confidence and interest in education. Others just want to find a job. Some of them find one during or after ABE courses. Others decide tot enter adult vocational education to improve their chances. Those who actually transferred to adult vocational and adult general education sometimes meet severe problems, like language problems, difficulties with the high level, time problems.
Improving the cooperation between ABE institutes and institutes for further education could be a solution to some of the problems. Some of the ABE institutes have high expectations of transition courses that combine training basic skills with vocational education and training (Moerkamp & Onstenk, 1991).

References

CBS, (1991). Volwasseneducatie, basiseducatie 1988/'89. 's Gravenhage.

Moerkamp, T., Emmelot, Y., Felix, C., Kat, E. de, (1991). Van basiseducatie naar een vervolgop-leiding. Eindrapport van het onderzoek naar de doorstroom van cursisten uit de basiseducatie naar vervolgopleidingen. Amsterdam, SCO, (in press).

Moerkamp, T., & Onstenk, J. (1991). Het leren van laag opgeleide werknemers. In: Basiseducatie, p.26-29.

TRAINING FOR EMPLOYEES WITH LOW EDUCATIONAL QUALIFICATIONS

J. Onstenk, Centre for Educational Research (SCO), University of Amsterdam, Amsterdam, The Netherlands

Introduction

In the paper results are given from a research project, commissioned by SVO and the Dutch Department of Education and Science, which reviewed participation of older workers with low educational qualifications (MAVO and less) in company training, described a number of specific training projects for these group of employees and analyzed motives and hindrances these employees have when participating in (formal) training. Also personnel managers and trainers were interviewed on their experience and evaluation of training for (older) employees with low educational qualifications.

Participation in training

Participation in formal training by older, poorly educated workers is very low: only 2 to 4% of all participants in adult education (whereas they constitute 17% of the working population). This is a very low figure, even if we take into account that training statistics are very unreliable and incomplete. Informal training activities (training on the job) are not accounted for in Dutch training statistics. Research from the United States (Carnevale a.o., 1990) and Germany (Weiss, 1990) shows that employees with low formal schooling receive mostly training-on-the-job, at the beginning of their jobs. Older employees however receive very little training-on-the-job.

Descriptive research shows that training for poorly educated employees, when they do receive any training, consists mainly of company specific, not certificated short courses, which contribute little to career possibilities (Van den Berg en Warmerdam, 1987).

Training is considered to be a hot topic in Human Resource Management and new production concepts, but until very recently poorly educated workers were largely excluded. This is partly caused by ascribed characteristics of employees (lack of motivation or learning capability), which lead managers to consider training not suitable for workers with low educational qualifications.

Previous research concluded that the most important reason for low participation in training is lack of possibilities of access to appropriate training courses (Wong en Siegerist, 1989). Training and learning in the company is not dependent mainly on individual motivation and preferences, but on availability of training opportunities. For highly skilled labor, professionals and management these in many ways are established and taken-for-granted. For not or half skilled labor these opportunities are lacking or just starting.

Changing training policies

A growing number of firms considers low participation of rank-and-file production workers to be a problem. They become aware that training needs to be integrated in global strategies of human resource management in the company. They develop training courses which are adapted to poorly educated workers, both in content and pedagogic structuring. On branch level central wage agreements of unions and employers provide funds for developing and offering training for these group.

In the research we interviewed personnel managers and workers in some of these firms. They are not representative for Dutch companies in general, because we looked for companies with an active training policy for poorly educated employees. But they share with a lot of firms the developments they encounter and which they try to deal with by training the work force. Skill levels are rising, because of changes in technology, organization, quality standards, growing customer orientation etc. The labor market offers no solution in meeting this demands, both quantitatively and qualitatively. So companies see a growing need for training their work force, including their employees with low educational qualifications. This need is met in different forms. The industrial firms we investigated develop specific training courses for production workers. Form and content of training courses are strongly practice and company oriented. Training is not only aiming at enhancing technical knowledge and skills, but to a remarkable degree also on social-

1092

communicative skills and on motivation and responsibility. This illustrates the fundamental change in expectations with regard to production workers. This change is not always taken into account by managers when discussing workers attitudes to training. As some workers complained, they are now expected to do what until quite recently they were explicitly forbidden to do by management. Pedagogically courses are designed to cope with learning and motivational problems which poorly educated workers might have. Courses are modular and include concrete examples and practice related tasks. Training content has direct relevance for the actual job. Experienced workers and lower managers are deployed as instructors.
The investigated companies in the administrative and caring services show a more traditional training policy. There are hardly any specific courses for poorly educated employees, but they are offered to attend regular courses, which they reluctantly do. Caring services (homes for the elderly) create new jobs at the bottom of the occupational hierarchy and offer training to attract and hold new employees, especially unemployed, older, poorly educated women, for these jobs. Clerical jobs have been upgraded or disappeared. Older poorly educated people hardly work in this sector anymore, apart from marginal administrative jobs for 'organizational left-overs'.
Evaluation reports show good results with regard to willingness and abilities of poorly educated workers to be successfully trained. This was also stated by the interviewed personnel managers and trainers.

Motivation and experiences of workers
In the research project some 100 workers were interviewed, of whom most recently had received training, for the first time in a very long time. Most respondents indicated that they were quite willing to participate in training. They showed positive attitudes towards training, which was oriented directly to the actual or expected job content. They evaluated recent training experience positively, when and because it was connected to improvement of their actual or future work content and competence. They preferred practical teaching forms. Experience with this specific training forms alters in some cases also general attitudes towards training. This positive attitude is conditioned by content and structuring of training which must be strongly connected with actual and prospective jobs. Workers have a positive, but realistic attitude towards training. They see good use for training, even if they do not see any real career possibilities or higher earnings opening up by training.
Learning attitudes of employees seem first and foremost to be oriented on functional and company specific training (knowledge and skills they can use in their actual job). This does not mean however that they have (only) an instrumental motivation. They strongly expressed their interest in training contents, because it opens opportunities to develop competence, to have better insights in the production process and backgrounds, to know better what they were doing, to improve quality etc.
On the other hand however some employees have mixed feelings on training policies of their companies. The most important point, especially in industry, is the obligatory character of training.

Many employees stress their right of choice, even if they feel strong pressure, because training is made conditional for further employment. Especially employees older than 50 made comment on this point. This is not only caused by the short period they expect to be in the company before retirement. It also is caused by their feeling that formal training policies take insufficiently into account their competence gained by work experience.

Production workers in industry show active learning attitudes, but a more passive attitude to training. They show great interest in learning new knowledge and skills to handle new machines, materials, products or quality norms, but they prefer learning on the job. Some respondents expressed strong concerns for devaluation of these way of learning. This does not mean however that they do not appreciate task oriented training opportunities. But they do not take active initiatives with regard to training. Sometimes this changes after having had some positive training experience. In some cases a wish to continue training for certification as skilled worker was expressed.

Respondents with the lowest educational qualifications (only primary school) express more than others fears with regard to being able to follow training successfully and to possible gains by training. This does not mean however that they evaluate more negatively training experiences they actually got.

Women, who restarted working in the caring sector, expressed much more an active orientation towards training, both for better doing the job and opening career possibilities.

Concluding remarks It can be concluded that training of workers with low educational qualifications, aimed at handling new technology or new organizational and qualitative demands can be successful, if it is integrated in company policy (including training facilities and involvement of line management) and is well designed pedagogically, taking into account practical and situated ways of learning of poorly educated adults.

Development of suitable materials and course models should be a major target for educational field in the coming years. More specifically attempts should be concentrated on developing suitable materials and assisting in designing training by company officials instead of (only) offering training courses on a contract basis. There is a growing need for training material which helps structuring training on the job (Kruijd, 1990). Next to that training the trainers with regard to learning of poorly educated people remains an absolute priority. This includes taking into account workers preferences for practical learning by experience on the job and by mutual help and explication by experienced workers and bosses (Thijssen 1988; Onstenk, forthcoming).

Training policies (including informal learning) should be integrated in strategic company policy. Training should not only be promoted by higher management. Especially with regard to training of poorly educated workers in the lower ranks of the organization active involvement of line management is crucial. Most specific training projects for lower educated workers are oriented to short term goals. Efforts should be made to extent opportunities in opening up training trajectories, which lead to certificates with external labor market value and skilled careers.

1094

References

Berg, J. v.d., & Warmerdam, J. (1987). Opleidingsactiviteiten in arbeidsorganisaties. Den Haag: Ministerie van Sociale Zaken en Werkgelegenheid,

Carnevale, A.P., Gainer, L.J., & Villet, J. (1990). Training in America. The organization and strategic role of training. San Francisco, Jossey-Bass.

Kruijd, D.(ed.) (1990). Opleiden op de werkplek. Deventer: Kluwer.

Onstenk, J.H.A.M. Leren en opleiden op de werkplek. Bunnik: RVE, forthcoming.

Thijssen, J. (1988). Het leren van de verouderende werknemer in veranderende organisaties. In: Pieters, J.M. (ed.), Leren en instructie in arbeidsorganisaties. Enschede: OCTO.

Weiss, R. (1990). Die 26-Mrd Investition - Kosten und Strukturen betrieblicher Weiterbildung. Köln: DIV.

Wong R., & Siegerist, E. (1989). Schoolbaarheid. Den Haag: Organisatie voor Strategisch Arbeids- marktonderzoek.

ADULTS BASIC EDUCATION, PARTICIPANTS AND COURSES [1)]

Ruijs, G.P.C.M.[1)], Netherlands Central Bureau of Statistics (CBS), Voorburg, The Netherlands

Introduction

Adult basic education covers activities for adults to acquire the basic adult knowledge, attitudes and capabilities necessary to function in private and public life.

Until 1987 training was mainly provided by volunteers, who had their roots in a wide range of target groups: women, ethnic minorities, etc. This background still has its impact on the access and dispersal of facilities to meet basic learning needs, as we shall see later on.

In 1987 the study facilities which had been covered by a number of educational regulations concerning open school, young workers' educational institutes, education for cultural minorities and literacy programmes, were incorporated in the Government Rule Basic Education.

Since then training adults in basic learning needs has become professionalized. Many institutes were merged into larger organization, and the course contents of the activities were mixed into an integrated curriculum of language, arithmetic and social skills.

---

1) The views expressed by the author in this paper do not necessarily reflect the official position of the Netherlands Central Bureau of Statistics

Participants

The most recent available CBS-data reveal that the target groups of the precursors of adult basic education are still well represented in the population of the participants today (text-table 1). The number of people participating have increased enormously.

As we can see in text-table 1, the main part of the participants' population consists of native women. After the first year of practice under Government Rule their share has reduced in favour of the foreign and non-native participants, who are now 50% of the participants,' population.

In 1990 over 1% of the adult population of the Netherlands participated in adult basic education; as for non-native inhabitans these figures are 3.5% for the male adults and 4.8% for the female adults.

Text-table 1

Participants by ethnic origin and sex

survey date	1987/'88	1988/'89	1989	1990
total[1)]	65 000	73 000	98 000	115 000
%				
native paticipants	58	51	52	50
men	16	15	15	14
women	42	36	37	36
non-native participants	42	49	48	50
men	18	21	20	21
women	24	28	28	28

Source: 1987/'88 survey by SVE; the remaining surveys by CBS.

1) The first two surveys refer to school years, the remaining two refer to calender years.

Activities

The course content of the activities can be described by four main aims:

- introductory courses (meant to ease further access);
- literacy programmes;
- programmes at the medium level of adult basic education;
- programmes oriented at outflow with high social participation (labour market, second level education).

The most recent data refer to the year 1990. Extra public finances provided for additional literacy programmes -fitting in the 1990 International Literacy Year- and additional outflow oriented programmes.

The impact of these campaigns are immediately reflected by the data. Comparing the surveys of the last two years, which both refer to calender years, we can see that the literacy programmes conveyed the strongest increase: the number of

1096

contact hours in literacy programmes in 1990 accelerated 1.8 times the number in 1989. This figure beats the acceleration of the outflow-oriented programmes (1.3), the introductory courses (1.2) and the medium-level courses (1.1) by far.
Till 1990 about 20% of the contact hours were spent on literacy programmes, half of which were consumed by native participants. As the Literacy campaign mainly focussed on the non-native participants, in 1990 the programmes for native participants are less than one-third of the total number of hours spent on literacy progammes.

Text-table 2
Activities by language offered and main aim

	1989	1990	1989	1990	
	abs. * 1 000 pch		%		1989 = 100
Total	7 178	9 234	100	100	129
Language offered					
Dutch (for natives)	2 582	2 709	36	29	105
Dutch (non-natives)	2 945	4 182	41	45	142
inc other than Dutch	880	993	12	11	113
unknown/irrelevant	771	1 350	11	15	175
main aim					
Introductory courses	380	454	5	5	119
Literacy programmes	1 382	2 547	19	28	184
Dutch (for natives)	695	734	10	8	106
Dutch (non-natives)	577	1 431	8	15	248
Medium level courses	4 127	4 565	57	49	111
Dutch (for natives)	1 241	1 363	17	15	110
Dutch (non-natives)	1 946	1 888	27	20	97
Outflow-oriented programmes	1 289	1 668	18	18	129
Dutch (for natives)	453	461	6	5	102
Dutch (non-natives)	373	732	5	8	196

N.B. pch = (number of) participants' contact hours

The additional programmes deriving from the Literacy campaign, as well as those deriving from the national Project on Additional Training, mainly benefitted the non-native participants. Text-table 2 shows that the number of contact hours in courses with Dutch as a native language increased by 5%, whereas those in

courses with Dutch as a non-native language increased by over 40%, a pattern that all the main aims have in common, except for the courses at the medium level. Most of the increased share of contact hours traceable to non-natives in the literacy programmes and the outflow-oriented programmes are chargeable upon the decreased share of non-native participation in the medium-level programmes.

EDUCATIONAL POLICY AND THE (NON)PARTICIPATION OF LOWER EDUCATED ADULTS IN ADULT EDUCATION

T.F. Brandsma, Centre for Applied Research in Education (OCTO), University of Twente, Enschede, The Netherlands

During the late seventies and early eighties, educational policy in The Netherlands with regard to adult education was characterized by an educational political and social cultural orientation. Adult education was perceived as 'second chance education', with an emphasis on lower educated adults. Which was, amongst others, expressed in the attention for the provision of educational programs and/or policy measures tailored to specific target groups, such as literacy programs, Open School-programs and the possibilities to participate in part-time secondary education during daytime or to enroll in only a part of the secondary educational program (taking only some subjects instead of the whole course).
During the first half of the eighties the orientation underlying the adult educational policy changed. Under the pression of labour market problems and a stagnating economy, more and more the contribution of adult education to socio-economic goals was stressed and the attention for continuing vocational training increased. Although the provision of both general and vocational programs tailored to lower educated adults still continued, this no longer was an aim in itself but one of the measures subject to the socio-economic goals.
Despite all the efforts during the last decades to increase their participation, the lower educated still participate (relatively) little in adult education. This is not only the case in The Netherlands, but in other European countries and the United States as well. This paper considers the causes and explanations for the low participation of lower educated adults in adult education:

- Which factors and mechanisms influence the limited participation of lower educated adults in adult education?;
- Which consequences can be derived from insight in the working of these factors and consequences for educational policy with regard to adult education in general and the participation of lower educated in particular?

The results presented in this paper are based on a review study, conducted in 1991 (commissioned by SVO and the Ministry of Education and Science), that considered the functioning of educational and training markets in the Netherlands. A major finding of the review study is that sound statistical data on participation in adult education by level of educational attainment are lacking in The Netherlands. As far as data are available, they are derived from seperate research projects focusing on participation in specific forms of adult education (basic education, secondary general education, vocational and training programs). Eventhough the

1098

picture is scattered, the findings of these different studies all point in the same direction; the number of lower educated adults participating in adult education is limited. The participation in special programs aiming at the lower educated shows a tendency of `creaming' and the rate of succes is rather low, especially in vocational training programs for lower educated, which are provided in the context of labour market policy. A substantial percentage of the participants drops out of the training scheme without achieving a certificate. Four explanations for the (non)participation of lower educated adults will be discussed in this paper: (1) shortcomings in educational policy regarding adult education; (2) characteristics of the provided educational programs in relation with (3) the lower educated adults' needs for and deterrents to participation and (4) the segmentation and non-transparency of the educational and training markets. It appears that these explanations are not as much conflicting, but complementary. Out of this the consequences for educational policy with regard to adult education (for the lower educated) will be discussed in reference to (more or less succesful) experiences and measures in other European countries.

1099

CHAPTER 8

HIGHER EDUCATION

1100

SEX DIFFERENCES IN THE LABOUR MARKET POSITION OF YOUNG GRADUATES

B.J.H. Lodder, P.J.E. van de Loo, G.W.M. Ramaekers, & R.K.W. van der Velden, Research Centre for Education and the Labour Market (ROA), University of Limburg, Maastricht, The Netherlands

Introduction

Commissioned by the Board of Governors of the University of Limburg (RL), the Research Centre for Education and the Labour Market (ROA) carried out a research project on sex differences in the labour market position of students of the RL, who graduated in medicine, health sciences, Dutch law and economics in the years 1986 up to 1989. The analyses were executed on data generated by the so called 'labour market scanner' which ROA developed for the RL. With this scanner the labour market position and the careers of the graduates of the RL can be periodically mapped out.

In this research project sex differences in the labour market position are examined under two headings:

- Sex differences with respect to labour market participation. Under this heading the following aspects of the labour market position are analysed: present participation, duration of unemployment before entering the labour market, number of working hours per week and job security.
- Sex differences regarding the match between education and job. Under this heading the following aspects are examined: occupation, required educational level, management position, income and the potential job mobility.

It is examined whether male and female graduates of the RL score significantly different on a number of aspects of the labour market position and whether these different scores result from the sex of the graduates or from intervening variables, such as age, education or job characteristics. At some points the results with respect to the graduates of the RL are compared with national figures.

Sex and choice of study

From national data one can conclude that the drawback of women in educational level is still growing smaller. Although men and women are still choosing different studies one can see that also this gap is more narrow than some years ago. At the beginning of the nineties more women than men are graduating in medicine. It is expected that at the end of this century as many women as men will graduate as

lawyers. The number of women graduating in economic sciences grows very rapidly, but the percentage share of women will be only 25 in the year 2000. On the contrary health sciences is and will be a study which is dominated by women; at present more or less 70% of the graduates are women.
Since the RL is a very young university and the number of graduates is still limited one can't already conclude whether the percentage share of female graduates is higher at the RL than elsewhere in the Netherlands.

Labour market participation

Rate of participation

Considering the rate of participation (the share of the labour force in the total population between 15 and 65 years), one can see from national data that in 1990 between 91 and 96 percent of the men who graduated in medicine, economics or law joined the labour force. For women these figures were in 1990 between 75 and 88 percent.
The differences between male and female graduates of the RL are evidently less. For the graduates in medicine, Dutch law and health sciences the rate of participation varies from 88% for women to 94% for men.

Unemployment

Although at the national level female academic graduates suffer more from unemployment than their male counterparts, this is not the case for graduates in medicine, law and economics. From the results of a proportional hazards model, used to estimate the duration of the first unemployment period of the graduates of the RL, it also appears that there are no differences between both sexes. Male and female graduates of the RL have an equal chance to get out of unemployment.

Part–time work

From the bivariate analyses of the RL–data it appears that men significantly more often than women have a full–time job. This is not the result of some intervening variables, since the multivariate analysis shows that sex is one of the factors that significantly influence the chance that graduates have a full–time or part–time job. This is a national phenomenon. In 1990 more than 20% of the working female academics worked less than 20 hours per week (men only 6%).

Job security

Finally the differences between men and women in labour market participation are investigated by considering possible sex discrepancies in the formal duration of labour agreements. For the graduates of the RL it turns out that men more often than women have a permanent appointment (25% versus 20%). At the national level this discrepancy is smaller: 92 versus 89 percent for all working men and women, and 91 versus 90 percent for working men and women between 25 and 29 years. From the multivariate analysis it appears that sex has no significant direct effect on the differences in having a permanent appointment or not.

Match between education and job

Occupation

Both national figures and RL data indicate that women are partly found in other occupations than men and that they are concentrated in a smaller number of occupations. The national figures also indicate that women more often than men have in occupations which benefit less from growing employment.

Required educational level

Almost one–fifth of the graduates with paid work performs a function below their academic level. As regards the required educational level, the analysis shows that men and women have equal chances of working at an academic level or not.

Management and income

It can be seen that men are significantly more often than women working in management positions. Of the men 36% had a management function, of the women 'only' 22%. Those graduates that have a salaried job were asked for their monthly gross income for full–time work. From the regression analysis it appears that men earn some Dfl. 270 per month, gross, more than women.

Potential job mobility

Among this population of young graduates, the potential job mobility is rather high. Almost one–third (29%) of the paid workers is looking for another job. We assumed that the likelihood of wanting to change jobs is higher for women because women, as a result of unequal labour market opportunities can be forced to accept jobs that are not their first choice and are therefore more likely than men to be looking for another job. However sex could not be shown to have any significant effect on the propensity to change jobs.

Conclusion

We observed that men significantly more than women have management positions, full–time jobs and higher incomes. These differences between male and female graduates can not be attributed to intervening variables, since the multivariate analyses indicate that sex has a significant direct effect on these three aspects of the labour market position.

These sex discrepancies can often be observed in labour market studies, which however usually concern the entire labour force and hence persons with different educational levels and in different stages of their working life. The differences are therefore often attributed to the fact that women interrupt their careers for several years to take care of young children. This explanation however is not plausible for the differences between male and female graduates of the RL, for they are just standing at the start of their careers.

The question remains whether the observed differences between the two sexes result from self selection or from discrimination of women in the labour market. The fact that women more often have part–time jobs seems to be a consequence of self selection. This can also apply to the fact that women are working less often in management positions. But the fact that women, although the data are converted to full time work and corrected for other job characteristics, have lower incomes than men seems to result from discrimination.

1103

PATHS OF LEARNING AND PROFESSIONAL INTEGRATION WOMEN AND THEIR COLLEAGUES IN ARCHITECTURE

U. Martiny, Hochschule fur bildende Kunste Hamburg, Germany

Introduction

Paths of learning and paths of professional integration might differ in significant ways for women und their male colleagues. This phenomenon is decidedly to be seen along social constructed forms of paths. Individual transfers of acquired knowledge and skills to new knowledge and task domains follow social constructs of paths. Paths and the steps of paths open or close and social chances inherent to higher education and to the structurized market. The empirical variety and predominance of certain forms of paths of learning within the core curriculum are put in a perspective with later paths of professional integration and marginalization and with leading ideas inherent in aims of higher education of architecture and regional planning.

Theoretical framework

Our approach works on the following assumptions:

- The construction of paths of learning and of professional integration is a gendered construction. The question of deconstruction - a diminishing of gendered constructs - and its maybe moving factors needs as its antecedent the analysis of gendered institutions. Within life course perspective it cannot be helped to at first get aware of overt and covert gendered situational and structural obstacles to identify the ways persons are able to assert themselves against those obstacles. Processes of mutual assessment between institutions of higher education and specific market conditions for the exercising of the profession show their counterpart on the micro level in the social construction of paths of learning and of professional integration. The leading ideas inherent in aims of higher education in the field of architecture and regional planning are object of mutual assessment between higher education and its personnel on one side and employers and customers on the other. Professional organisations of architects and regional planners may intervene this interplay with their leading ideas of what higher education in this field should aim at.

Special interest is given to social deconstruction of gender and its theoretical as well as empirical prerequisites. This includes the methodological and political implications of treating women as exceptions from the rule.

- Defining paths of learning and professional integration for remarkable different groups (e.g. of age groups, of types of institution of higher education, of regional labour markets in federal states and of gender inequality in social chances to assert oneself) follows ideas of grounded theory. This means: a) empirical generalisations on processes and the social construction of processes are put forward. They get formulated removed from universalistic concepts. Universalistic concepts are not appropriate because they tend to ignore the construction of gender as one of the sources of social asymmetry. Another affinity to grounded theory b) takes up the fact that the methodology of life course research is still in an explorative stage. Concepts get developed during the research process itself and cannot be taken over from the heritage of welldocumented established theorizing. The same applies to studies on the social construction of gender throughout the life course and of more or less specifically gendered paths of learning and professional integration.

Modes of Inquiry

To investigate possible indicators showing an improvement of social symmetry between men and women in professional appearance and in long-term effects of the impact and outcomes of higher education our concepts and methods are oriented on the macro and micro level of analysis and their respective links.
Four lines of investigation are followed:

- inquiring the direction of qualification in institution of higher education (in their regional mixture), its aims and inherent ideals or models;
- investigating market conditions in the professional field, market selection, needs of a regional structured market and met an unmet demands;
- inquiring the respective fitting with institutional conditions of the other sphere: Mutual ideals or models of higher education and corresponding ways of selecting, qualifying and gratifying in the spheres of higher education and on the part of employers and customers;
- To clarify the social construction of individual paths of learning and professional integration we finally link results of macro analysis with an analysis on the micro level: qualification strategies of women and their male colleagues, the planning of a professional career, self-perception and professional appearance which depends on these factors are being compared.

Remarkable combinations of these factors are put in a perspective with institutional ideas or models of higher education and with the social reality of specific gendered steps and paths, built on these ideas.
Speaking of methods: there exists a model of generating data and combining different sources of data which has been developed and revised in the study s first phase. The main source of information comes form socio-biograhical interviews. The interviews consider remarkably different employment and work situations as an initial point.
The text analysis of interviews is completed by three other methods:

- an analysis of institutionally designed steps and paths of study and of its underlying aims and inherent ideals or models of higher education as they appear in higher education, in the structurized market and in professional organization;

- a gathering of process data on educational achievement and specialisation as well as of - conventionally speaking - a more or less successful labour market integration (using archives of "Architektenkammern" that is of professional organizations practicing some self directed policies of cooptating and further education e.g.);
- a secondary analysis of mass data on higher education, together with a gathering of unpublished mass data from universities, technical colleges, art colleges and "Fachhochschulen" (which might be compared with polytechnics although there ist no direct equivalent in other European higher education).
Several northern German federal states will be compared (Hamburg, Nordrhein-Westfalen, Hessen, Niedersachsen) with their regional mixture of core curricula in the field of architecture and regional planning and their regional market conditions and forms of professional self-organization. Different age groups of men and women will be distinguished.

Discussion of first results

The construction and deconstruction of gendered paths of learning and professional integration and its maybe moving factors are the focal points of the paper. Situational and structural obstacles for social symmetry between women and their male colleagues in architecture are being analyzed. The investigation centres on higher education and its manifold social definitions. Corresponding but different inherent leading ideas of the aims of education are confronted, linking macro and micro level, in examing questions of gendered construction of these ideas and of social reality built on these ideas. There are also some empirical results presented for one of the regions (Hamburg) and an outline of the model which will be used for all northern federal states of Germany included in the study.

DISCIPLINARY CULTURES

Tony Becher, University of Sussex, Sussex, United Kingdom

In reviewing the significance of disciplinary cultures for an understanding of academia - and *a fortiori* as a framework for research in higher education - it would seem appropriate to begin with a brief discussion of the concept of culture and a reminder that disciplines can, for the purpose of analysis, be regarded as comprising both a cognitive and a social aspect. It is with the latter that the subsequent discussion will be primarily concerned.

My own recent research has involved a study of cultural norms and practices in twelve contrasting disciplinary fields (physics, chemistry, biology, mathematics, geography, economics, sociology, history, modern languages, law and engineering) and a more specific enquiry into graduate education in six of these. It has been useful in the context of these investigations to differentiate between three cognitive levels, namely broad knowledge areas (which can be categorised as hard pure, hard applied, soft pure and soft applied), individual disciplines, and sub-disciplinary specialisms. The focus in this paper will be on the second, disciplinary level, though the other two are also important in a number of respects.

Disciplinary cultures, it can be argued, pervade all strata of higher-education, from the macro to the micro. To a significant extent, they cross not only institutional but national boundaries, and are thus also of relevance in comparative research. The evidence for this claim stems from the findings reported in my *Academic Tribes and Territories* (1989), which related to both British and American universities, and from the more recent exploration of graduate education, which was one element in a five-country study embracing the UK, France, Germany, Japan and the USA (to be published as Clark B R (ed) *The Research Foundations of Graduate Education (1993)*).

The significance of the theme at the macro level relates - to take one example - to the general forms of interaction between higher education and the wider society. Here, there are evident differences between professionally-oriented disciplines such as engineering and law at the one extreme and literary, linguistic and historical studies at the the other: but there are also notable variations in pattern within the pure sciences as well as between them and the humanities. Other significant dimensions at the macro level include questions of access (since admissions criteria can be shown to vary significantly from one subject area to another in systems where access is not totally unrestricted); the social implications of differential gender preferences between subjects; the labour market potentialities of particular types of disciplinary qualification; and the opportunities to engage in contract research and sponsored vocational training.

1107

At the meso level, disciplinary differences are a major consideration in institutional management, where distinctions between labour-intensive and capital-intensive fields are only among the most evident issues for consideration. Such distinctions affect both the allocation of resources and the evaluation of performance. Given the multiplicity of variations between one subject and another, it is demonstrable that no single set of performance indicators can be fairly applied across the range of departments within any non-monotechnic institution.

It is also possible to identify a number of noteworthy issues relating to the micro level of the individual department and its associated activities. Among the themes to be explored here are major differences in research modes and the related patterns of graduate education; contrasts in approaches to undergraduate education; consequent distinctions in the requirements for staff development and the teaching of study skills; the need for alternative practices in curriculum design; and variations in the patterns of departmental management.

The final issue to be explored will be why so much of current higher education research apparently ignores or sidelines these types of cultural contrast.

PROCEDURE TO PROMOTE EFFECTIVE AND EFFICIENT STUDY SKILLS (PROPES)

Joke Oosterhuis-Geers, Educational Centre , University of Twente, Enschede, The Netherlands

Introduction

In 1989 the faculty board of the Department of Public Administration of the Twente University was concerned about the low passing rates of their first year students. The percentage of students that passed the number of exams necessary to graduate within four years, was about 30% instead of the planned 70%. Because of this the faculty board requested the Educational Centre to organise training in study skills. But low passing rates are not necessarily due to poor study skills, they might also be caused by poor teaching skills, an imbalance between timetables for different courses and unrealistically high standards of achievement etc..

Therefore some lecturers and students were interviewed about those aspects. Also all first year students of 1989/1990 received a questionnaire about their study habits. Based on these data we concluded that some courses had to be improved, but also that some students were in need of better study strategies. Our survey also showed that students were not very willing to change their study habits. Most have just one solution to accelerate their low study pace: 'to work harder'. Even students, who spend a lot of time studying kept saying that they had to spend more time in studying.

How to change this mind set of student? We decided to perform a literature search on this subject.

Learning to learn implies two important learner characteristics: intelligence and motivation. Educational and psychological research has been developed in two directions: (meta-)cognitive theories (like information processing theories) and motivational theories (attribution theory) (Pintrich a.o.,1986; Weinstein & Mayer,1986).

Procedure to promote effective and efficient study skills

To integrate both approaches of study skills we developed a procedure to promote effective and efficient study skills (PROPES).

This procedure consists of 5 steps.

The first step involves a *provisional advice*. The Dutch government has provided by law that faculty boards give their first year students an advice at the end of their first year study. This advice concerns their ability to go on with the study or to leave and look out for other opportunities. The first step of PROPES stimulates the

1109

process of students' reflecting on their own study pace and motivation by giving a provisional advice immediately after the first trimester.

The second step involves the confrontation of first year students with experiences and ideas of fellow students from the second year of study. This step involves showing a *videotape* on which 5 interviews with fellow students are recorded. In these interviews each student tells his or her own story. All those stories have four aspects in common: the experience of failing an exam, the search for a cause, the improvement of study skills and the subsequent success in the examination. The reasons for this step are as follows. Many students attribute their low passing rates to one cause: not working hard enough. In many cases attributing exam failure to this cause is a mistake. By showing students that the causes of failure are quite differentiated and that they can be changed, the students become now aware of possible causes of their own underachievement. Due to this awareness some students might change their study habits on their own. Other students might need more support to change their behaviour. Therefore we developed the following step.

The third step involves giving students the opportunity to *diagnose* their strengths and weaknesses in study skills by means of a questionnaire. After answering the items in the questionnaire the first year students can count their scores themselves and look at the advice that belong to certain scores. The advice varies from references to parts of a book about study methods to training in study skills. The reason for this step in the procedure is twofold. In the first instance we assume that students differ in their mastery of several skills. Training in study skills should take into account these differences, because it is a needless waste of time to teach students skills they have already mastered. The other reason is that some students need only advice and some reference to a book, while others really need to be trained.

The fourth step involves the *training in different study skills*. Only those students who were advised about training in a certain skill, are allowed to participate in that particular course.

The fifth step is about the *involvement of mentors* who supervise first year students. They have to be informed about this procedure, so students can ask them for support if things don't go as planned.

By means of this procedure first year students are confronted quite early with their study pace, so they might reflect on this and their motivation. As a consequence they become conscious of study problems and possible causes. We expect this will result in more adequate attributions of their achievements. But adequate attributions are not always enough to change study skills. The procedure provides means for specifying attributions by developing a questionnaire and for behavioural change by giving advice and offering opportunities for training in study skills.

The main goal of PROPES is making students more responsible for their own learning process. This also implies that each step screens students who continue with the next step or continue on their own. I call this the 'screening function' of PROPES.

1110

Research results

Research on the effects of PROPES concentrates on effects on the micro and meso level. Research questions are:

- Are the goals of step 1 to 5 achieved ?
- What are the effects of PROPES on study pace and study results ?
- Does PROPES perform the 'screening function' ?

To find an answer to question 1 and 3 we made use of the quasi-experimental 'untreated control group design with pretest and posttest'. To answer the second research question we used the 'interrupted time-series design with a non-equivalent no treatment control group'.

In the academic year 1990/1991 the PROPES was partly implemented due to practical circumstances. The videotape was not ready at that time.

In the months December and January 1991/1992 the PROPES was fully implemented in two faculties of the Twenty University (Faculty 'Public Administration' and 'Informatics').

Formative evaluation of the partly implemented PROPES gave some evidence that:

- Step 1 provoked a defensive attitude, because the provisional advice was perceived by the students to be premature;
- Step 2 was not implemented;
- Step 3 achieved the goals; all first year students answered the questionnaire;
- Step 4 was partly achieved; only a few students were trained in study skills; most other students did not change their study habits on their own;
- Step 5 achieved the goals;
- There was no effect on the study pace of the total group of first year students;
- Students who were trained in study skills improved their study skills and their study pace, and
- the screening function was realised in a different way than expected.

Because of the above mentioned result the steps of PROPES have been partly changed in 1991/1992. The main characteristic of the change concerns the involvement of the mentors in the implementation of step 1, 2 and 3.

The changes in the PROPES were as follows.

Instead of a written provisional advice each mentor of the faculty plans a meeting with his students. At this meeting the mentor confronts his students with the study pace of fellow students in former years by showing them some statistical information about study rates and study pace. After this short confrontation the mentor and his group of first year students look at the videotape and discuss the theme 'failing exams and locus of control'.

At the end of this meeting students were free to take home the questionnaire about study methods and send it back to the Educational Centre. Students who answered the questionnaire got the results and their mentor also.

After this the mentors invited these students to talk about the results. Based on these consults students could subscribe to a course in one, two or three study skills.

The first analyses of the effects of PROPES as it was implemented in 1991/1992 showed the following results.

1111

- Step 1 and 2 achieved their goals. Although most students still underestimated the possible risks of their low study pace, they discussed seriously the subject 'failing exams, locus of control and study methods';
- About 30% of the students answered the questionnaire about study methods;
- Excellent students (study pace > 100%) and non-motivated students (study pace < 15%) did not answer the questionnaire;
- Almost all students who consulted their mentor about the results of the questionnaire were convinced of being able to change their study habits by themselves;
- Because of this we decided to offer twice (after the first and after the second trimester) the opportunity of training in study skills;
- Results in terms of effects on study rates and study pace have not been gathered yet, but will be presented at the conference in June.

Point of discussion

An interesting point of discussion in June might be the following:

- Students seem not very willing to change their study habits. With PROPES we try to stimulate students' responsibility for their own learning process. But others propagate integration of study skill training into the obligatory part of the first year programme.

What about a better marketing of PROPES? How can we convince first year students of the benefits of study skill training? Statistical information, gathered by scientific research on study skills and study results seem to convince only the scientists themselves.

WHY SOME GIRLS DO DROP OUT OF TECHNICAL STUDIES AND OTHERS DON'T

Hetty Grunefeld, University of Twente, Enschede, The Netherlands

Introduction

A few years ago our university grew interested in the question why girls drop out of their study. It seemed that proportionally more girls dropped out than boys. In reality, the percentages of female and male drop-outs were not significantly different. The token theory of Moss Kanter works: females who drop out attract attention.

Research into this problem was started by the Educational Centre in two departments: Computer Science and Mechanical Engineering. We wanted to know whether there are differences between female students who drop out of their study and others who don't, and if so, what actions we could carry out to keep female students at our university. We were looking for improvements at several levels, being curriculum, course, teacher and student level.

Theoretical framework

In 1985 the Women's studies group of the University of Twente published the results of a survey on study careers of female students in Twente (Everts, 1985). They concluded that several processes effect study careers, but the effects were varying for every student. The most important processes they found were on the one hand social integration in the department and in the living environment and on the other hand the way female students handle conflicts with regard to self esteem and to time and energy management.

Another source we used is a literature study by Smit and Oosterhuis (Smit, 1988). Factors we included in our study based on (Smit, 1988) are self image, role conflicts and minority-position problems (token theory of R. Moss Kanter), hobbies, learning style, instructional media and course material.

Our hypothesis is that drop-out students have more problems associated with these factors than students who did not drop out.

Research method

Because of the exploratory character of the research we chose for depth interviews with a small number of respondents. We did not have male respondents, because we were interested specifically in problems of females.

We used three criteria in the Computer Science (CS) department:

1113

- equal numbers of drop-out and studying girls;
- same cohorts, (1984, 1985, 1986, 1987);
- approximately the same results in secondary school.

As a consequence of these choices, we had in de CS-department a group of students with nearly the same starting position. Therefore, the differences we would find were not explicable with different capacities, but with some of the factors as aforesaid.

In the department of Mechanical Engineering (ME) we had another strategy. Our aim there was to find as many reasons as possible for dropping out, or to continue with ME. Our group of respondents had, among other things:

- different secondary school results (the drop-outs) or study pace;
- lengths of stay at university;
- different cohorts, (1982 - 1989).

In the Computer Science department there were 29 respondents, 15 of them had dropped out. In the Mechanical Engineering department there were 25 respondents, 7 of them drop-outs.

About half of the interviews was held face to face, the other half consisted of written interviews. We found no reason to think that this would bias our results.

We sorted the answers in categories and counted them. Statistical tests were done with the Fischer Exact test (Everitt, 1979), with $p < 0.01$ or $p < 0.05$.

Results

We will describe a selection of the subjects of the study careers of our respondents.

*: statistically significant	C.S. Drop N=15	C.S. Still N=14		M.E. Drop N=7	M.E. Still N=18	
Study motivation						
CS/ME was first choice	15	10	*	3	8	
CS/ME was expected to be as laborious as sec. school	6	0	*	0	7	*
Next study or work is technical	10	-		6	-	
Social integration and position of women						
Kind of contacts between man and women tolerable	9	12		1	9	*
Number of female teachers in the department too low	2	7		7	5	*
Position of women in the department comparable to position of men	9	9		0	8	*
Future possibilities						
Wants to have a job and knows what & where	11	7		1	13	*
Wants to have children and a career	11	11		5	8	

CS-drop-outs get in the first trimester 37% good marks, the others 62%. ME-drop-outs get in their first year 52% good marks, the others 74%.

CS-students have the same preferences for and dislikes of specific courses, just like ME-students.

We asked whether they would like to participate in specific activities for female students. The answer was: "That depends on the programme."

Conclusions

The study careers of female students who dropped out and who continue studying, have many similarities. We found differences only in study results, expectation of difficulty and study methods. We did not find that lack of social integration is a problem that leads to dropping out. Apparently, the increased number of female students at our university made it easier to feel integrated.

The reason to stop is getting poor study results. We checked the results and indeed, we think they made the right decision. We can explain their poor results by looking at their study method and their expectations of the study.

Recommendations: try to improve the publicity and information about the study. However, there will always be some people who do not listen to all our information, so we must not expect too much of this recommendation.

In Computer Science we signalled a problem with prerequisites for the programming courses. Almost all respondents remarked that the start of these courses for students without computer experience was too fast.

Recommendation: introduce students in using of computers, give them an opportunity to gain experience with and understanding of computers. The CS department has implemented this recommendation, the new curriculum starts with such a course.

We asked the girls if they had suggestions for better course material. They had not. How should they know? Our recommendation is (based on (Smit, 1988)): experiment with feminine ways of education. We expect that fewer females will drop out, because their learning style will then harmonise with the education style.

In the department of ME, there is only one female teacher. Especially dropped out students ask for more female teachers. Recommendation: start a positive action programme.

We found that almost all of the female respondents expected that, if they want children and a job, they should have a part time job, because they think the father of the children wants to work full time. This means they prepare for another kind of career than most male students do. So, already during their study they put themselves behind men.

Recommendation: introduce life/career planning courses in the third/fourth year of the study, for both women and men. The ambition of such a course: emancipate men!

References

Blessing, L.T.M., Van Bruxvoort, J. ,Van den Kroonenberg M., & Oosterhuis-Geers, J.A. (1991). Blijven studeren of afhaken bij WB. Een onderzoek onder meisjes die Werktuigbouwkunde studeren of gestudeerd hebben. Doc 91-56, Enschede: Onderwijskundig Centrum, Universiteit Twente.

Everitt, B.S. (1979). The Analysis of Contingency Tables. London: Chapman and Hall.

Everts, S. & van Oost., E. (1985). Vrouwelijke studenten aan de TH-Twente. Een verkennend onderzoek naar de studieloopbaan van vrouwen in technische studierichtingen. Enschede:Technische Hogeschool Twente.

Grunefeld, H., van Petersen, G.M. (1991). Blijven studeren of afhaken. Een onderzoek onder meisjes die Informatica studeren of gestudeerd hebben. Doc 91-57, Enschede: Onderwijskundig Centrum, Universiteit Twente.

1115

Smit, N.J. & Oosterhuis, J. (1988). Samenvatting literatuuronderzoek vrouwenstudies. Doc 88-26, Enschede: Onderwijskundig Centrum, Universiteit Twente.

1072

19 i 6

EDUCATIONAL PROGRAMMING RELATED TO STUDY PROGRESS

Ellen P.W.A. Jansen, Centre for Research into Higher Education (COWOG), State University Groningen, Groningen, The Netherlands

Introduction

The last two decades the central government in the Netherlands aims to decrease the duration time of students at the universities. In 1982 a new law (Wet Tweefasenstructuur) was accepted that limited the duration of courses to reach a university degree, to four years (with some exceptions like medicine) and that limited the maximum number of years a student may spent at a university to six. Also universities were obliged to have in each field of study a propedeutic exam after one year study. The maximal number of years to reach this exam was limited to two years.This new law and some other developments in government policy forced universities to pay much more attention to monitor the study progress of students. One way to do this is taking measures on the level of educational programming. In accordance with these ideas the department of social sciences at the University of Groningen (the Netherlands) has changed in 1989 the principle of programming of courses for all the sections of the department. From that time each section (psychology, sociology and educational sciences) had to programme their courses in the same time period of 6 x 7 weeks. The purpose of this uniformity of planning was to improve the efficiency of the department. Efficiency was measured by study progress and exchange of subjects between the several sections of the department.

The problem of study reported here is to what extent de study progress of students has improved since the introduction of the principle of uniformity of programming, and to what extent students report to choose or not to choose subjects in another section than their field of study related to the programming in 6 x 7 weeks or not.

Theoretical framework

An important way to make university education more efficient- i.e. to improve the study progress of students- is to spread the study load. By a better spread of study load over the year students are able to use their time for studying in a most optimal way (Van der Drift & Vos, 1987; Jansen & Joostens, 1989; Knippenberg & Loos,1983). Research in different fields of study showed a relation between the spread of study load (by spreading the preliminary examinations) and the study progress of students (Zwaal, Joostens & Jansen, 1987; Welling & Jansen 1989;

1117

Knippenberg, 1989). By another way of educational programming a better spread of study load can be reached: shorter periods of instruction followed by conclusive tests. Students are forced by this measure to study more regular. When the period of instruction is long, for instance 5 months, students are in the habit of studying during that period on a low level and work only very hard at the end of that period when the preliminary examination is planned.
For psychology-students that are behind schedule Meerum Terwogt-Kouwenhoven (1990) found that the factor 'problems with using their time for study' was one of the important explanations for study success or failure. For other fields of study this was confirmed by de Jong, Koopman & Roeleveld (1991). Also at the University of Groningen in the field of economics problems with planning their time for study was reported as one of the main causes for drop out or getting behind schedule. So the introduction of the planning principle of 6 x 7 weeks seems to be one of the successful measures to improve the study progress of students. The other advantage of the fixed periods all over the department should be the facilitation of exchange of subjects from different fields of study. In the old system concurrence between courses existed, while fixed points in time where students could start with subjects were missing. This was possibly one of the reasons for students not to choose courses out of their own field of study.

Method
The study progress of students who started to study before 1989 is compared with the study progress of students who started in 1989 or 1990. Measures for study progress are:

- the percentage students that passed the propedeutic exam after 1 and 2 years;
- the realized study-laod on several points in time.

The data are analyzed on the one hand with ProgRESS, a programme for registration and evaluation of study progress developed by our institute and on the other hand with SPSS/PC. To find out why students choose or do not choose subjects out of their own field of study and which problems they see to do so, due to the old programming pinciple, we questioned students who started their study in 1986 and 1987. The choice for the cohorts 1986 and 1987 was made because students in the third and fourth year of study often are free to choose some subjects themselves.

Results
The field of psychology programmed their education since 1982 in six periods of seven weeks, the field of sociology since 1988 and the field of educational sciences since 1989. The field of educational sciences however did programme their education for the propedeutic exam in 1987 and 1988 in full blocks, i.e. one subject was instructed in a short period directly followed by a preliminary examination. Comparing the study progress in terms of numerical returns (i.e. the percentage students that pass the propedeutic exam in one or two years) in the semester system with the shorter period system (we make no distinction whether 6 x 7 or full blocks) we see a significant difference between the two systems in favour of the shorter period system (table 1).

Table 1
Difference in rate of numerical returns between two programming systems

	Pass their propedeutic exam	Stand. dev.	t-test
After 1 year			
short period	41	.492	p = .001
semester	29	.457	
After 2 years			
short period	79	.405	p = .008
semester	70	.458	

(source: I. de Vette, 1990)

So it seems that a shorter period of instruction followed by conclusive test is giving better results. If we compare within one field of study the number of months students need to reach the propedeutic exam, we see that for the field of educational sciences students need significantly less months in the system with full blocks respectively short periods; for sociology the results seem only better for cohort 1989 (table 2).
Also the realized study load (obtained number of creditpoints) is in general higher in the system with blocks or short periods.

Table 2
Average number of months (x) needed to reach the propedeutic exam

	x	t
Educational sciences		
coh 86 (semester)	16.5	
coh 87 (blocks)	12.6	p = .000
coh 86 (semester)	16.5	
coh 88 (blocks)	11.9	p = .000
coh 86 (semester)	16.6	
coh 89 (short per.)	13.6	p = .000
Sociology		
coh 87 (semester)	16.4	
coh 88 (short per.)	17.7	p = .383
coh 87 (semester)	16.4	
coh 89 (short per.)	13.0	p = .007

The survey on planning- and choice behaviour of third and fourth year students resulted in several problems for students for chosing subjects out of their own field of study. Those problems were partly caused by the different planning principles used by the fields of study. Time table problems were often mentioned, like two preliminary examinations on the same time or a full block during two or three weeks while also weekly instructions had to be followed.
However, mostly students say to choose subjects within their own field of study for reasons of interests in content or hope for a better position on the labour-market.

Because the students in this survey studied only partly under the short period system, conclusions on the subject of reaching the goal for another choice-behaviour cannot yet be made . In 1992/1993 a comparable survey will be held among students in their third, fourth and fifth year in the new short period system.

Discussion

We have seen that study progress can be influenced by changes in educational programming. We found that full blocks give even better results in study progress than short periods. Now the policy is to hold the department on strict rules for instance for the weeks in which preliminary examinations have to be hold. This however could have negative effects on the flexibility of programmes.

References

Drift, K.D.J.M. van der & P. Vos (1987). Anatomie van een leeromgeving.Een onderwijseconomische analyse van universitair onderwijs. Lisse:Swets & Zeitlinger.

Jansen, E.P.W.A. & Th. H. Joostens (1989). Blokonderwijs:Voordelen of vooroordelen.Onderzoek van Onderwijs,18, pp.46-48.

Jong, U. de; P. Koopman & J. Roeleveld (1991). Snelwegen en slingerpaden in en om het hoger onderwijs. 's-Gravenhage: Sdu Grafische Bedrijven.

Knippenberg, P.L.M.J. (1989). Evaluatie van de propedeuse Wiskunde en Informatica 1988-1989. Groningen: COWOG/Rijksuniversiteit.

Knippenberg, W.J.M. & F.A.J. Loos (1983). Doelmatig hoger onderwijs? Verhoging van het studierendement. Delft:: Delftse Universitaire Pers.

Meerum Terwogt-Kouwenhoven, K. (1990). Niet gewogen, toch te licht bevonden. Analyse van de rendementsproblematiek aan de Universiteit. Kampen: Mondiss.

Vette, I. de (1990). Evaluatie modulair onderwijs. Groningen: COWOG/ Rijksuniversiteit.

Welling, E. & E.P.W.A. Jansen (1989). Studieprogramma Kunstgeschiedenis. Studievoortgangsanalyse van de cohorten '83 t/m'87. Groningen: COWOG/ Rijksuniversiteit.

Zwaal, W., Th. H. Joostens & E.P.W.A. Jansen (1987). Onderwijsprogrammering en doorstroming van studenten. In J. Scheerens & W.G.R. Stoel (red.) Effektiviteit van onderwijsorganisaties. Lisse: Swets & Zeitlinger.

THE RELATION OF PERCEIVED SELF-EFFICACY AND GOALS TO ACADEMIC PERFORMANCE

A. Vrugt, M. Langereis, & Joh. Hoogstraten, Faculty of Psychology, University of Amsterdam, Amsterdam, The Netherlands

In several publications, Bandura (1989) has put forth and developed the concept of perceived self-efficacy. Perceived self-efficacy can be defined as an individual's estimate of his or her capability of performing a specific set of actions required to deal with task situations. Perceived self-efficacy is hypothesized to be an important determinant of action, and therefore, given the appropriate level of skill, of performance. According to Bandura, perceived self-efficacy is causally related to actions independently of the individual's actual demonstrated ability, even though the two are related. The same degree of actual success in performing a task can lead to different degrees of perceived self-efficacy, because individuals may reach different conclusions from the same perceived attainment. These different percepts of self-efficacy will be reflected in the subsequent performance of a test.
According to Bandura (1986) perceived self-efficacy is a cognitive factor that plays an influential role in the exercise of personal control over motivation. It is partly on the basis of self-beliefs of efficacy that people choose what challenges to undertake, how much effort to expend in the endeavour, and how long to persevere in the face of difficulties. That strong beliefs in one's efficacy heighten level of effort, perseverance and performance in difficult pursuits is corroborated by evidence across different domains of functioning varying from pain endurance to academic performance.
Wood and Locke (1987) have studied the relation of perceived self-efficacy and goals to academic performance. They developed a measure of academic self-efficacy (ASE). Wood and Locke distinguished seven relevant task areas: class concentration, memorization, exam concentration, explaining concepts, understanding, discriminating concepts and note taking. For each task area several items with equal intervals were developed to measure self-efficacy in relation to a course and course material. Consistent with Bandura's methodology, subjects were asked to indicate whether they could achieve the described level of performance (yes or no). The sum of the answers to these questions constituted the self-efficacy strength scale (SES). Subjects were also asked about their degree of confidence in their ability to perform at a described level (on a 0 to 100 scale). The sum of these scores constituted the self-efficacy magnitude scale (SEM). There were two to six items per task area and 29 items in all. All items were answered with respect to the course in question. Undergraduate students

1121

answered the questions while they were following the course before the exam(s). They also answered four questions about grade goals (e.g. which grade will you actually try for in the exam?). Subjects' responses on these items were averaged. In addition, the number of points in the relevant course, based on exams, was used as measure of task performance. Hierarchical regression analyses showed that, after controlling for ability and Grade Point Average, SEM and grade goals contribute significantly to task performance.

We have replicated the study of Wood and Locke, using the same procedures and measures and we will use the same analyses. Our subjects were undergraduate students in psychology who answered the questions in relation to three courses: introduction to psychology, social psychology and personality psychology.

People impose goal preferences on activities that reflect their basic orientations to achievement across a wide range of situations (Bandura, 1986; Elliot & Dweck, 1988). The question is how people's conceptions of ability affect the goals they pursue, which in turn determine the quality of their intellectual functioning. Two major conceptions are distinguished. In one perspective intelligence is construed as an incremental skill that can be continually enhanced by acquiring knowledge and perfecting one's competencies. Persons with this conception adopt learning goals. They seek challenging tasks providing opportunities to expand their knowledge and develop their competencies. Such an outlook will promote a mastery-oriented response to obstacles: strategy formulation and sustained performance. In the contrasting perspective intelligence is construed as a more or less stable entity. These people tend to favour performance goals. Quality of performance is regarded as diagnostic of intellectual capability. Errors and performance insufficiencies carry personal threat and may result in low ability attributions and a helpless response.

Our subjects answered six questions about learning and performance goals and about their mastery-orientation or helpless responses in relation to failure (twelve items). We used statements; subjects could indicate their agreement on bipolar six-point scales. We will examine whether these goals function as moderator or mediate between perceived self-efficacy and performance. For that purpose we will use analyses of variance and partial product moment correlation coëfficients.

References

Bandura, A. (1989). Self-regulation of motivation and action through internal standards and goal systems. In L.A. Pervin (Eds.) Goal concepts in personality and social psychology. Hillsdale, NJ: Lawrence Erlbaum.

Elliot, E.S., & Dweck, C.S. (1988). Goals: An approach to motivation and achievement. Journal of Personality and Social Psychology, 54, 5-12.

Wood, R.E., & Locke, E.A. (1987). The relation of self-efficacy and grade goals to academic performance. Educational and Psychological Measurement, 1013-1024.

CONSTRUCTIVE LEARNING IN HIGHER EDUCATION

J.D.H.M. Vermunt, Department of Psychology, Tilburg University, Tilburg, The Netherlands

In higher education a growing interest can be seen in instructional methods that promote constructive and discourage reproductive learning. Constructive learning is defined as learning in which students actively build their own knowledge by employing deep and selfregulated learning activities. The study described here was meant to gain insight into the thinking strategies students employ in a traditional university learning environment, as well as in the relations of these strategies with personal, contextual and study success variables. Our research is guided by a theory on the interplay between internal and external regulation of learning and the design of process-oriented instruction. Process-oriented instruction is defined as instruction directed at the initiating, supporting and influencing the thinking activities students use to learn and study. This theory can be placed within the constructivist movement that arose recently within educational psychology and instructional design. The research questions addressed were:

- What relationships do exist between students' use of processing and regulation activities, study orientations and conceptions of learning and studying?
- What are the relationships between students' learning styles and personal and contextual factors like subject area, type of learning environment (traditional versus distance), prior educational experience, age and sexe?
- How do students' learning styles relate to three indicators of study success: mean exam scores, percentage of exams passed, and study speed?

A diagnostic instrument, the *Inventory of Learning Styles*, that was originally developed for use within the distance education context of the Open university, was adapted for use within the more traditional 'lectured' education context of most other universities. This instrument was administered to 795 Tilburg University students of seven different subject areas. Also data on students' academic major, prior educational experience, age, sexe, exam participation and exam results were gathered. The data were analyzed via regression, factor, correlational and reliability analyses.

Concerning the first question above it turned out that the four learning styles originally identified with Open university students could clearly be recognized in this study. Moreover, the results showed that the strategies students used to process the subject matter were primarily influenced by their self-regulation strategies and much less by external sources like teachers or directions in the study materials. With regard to the second question, the data indicated that different learning styles seem to have different sources. The constructive learning

style was most associated with subject area, age, and type of learning environment. The reproductive learning style turned out to be most associated with prior educational experience and subject area. The applicative learning style was related most to subject area and type of learning environment. The problematic learning style was most associated with type of learning environment, and to a lesser degree with age, prior educational experience, and subject area.
Concerning the third question, the results showed that learning styles explained between 20 and 55% of the variance in study success. In general, the constructive style was positively associated with different indicators of study success, and the reproductive and problematic learning styles showed negative relations. The applicative style was rather neutral with regard to exam results. There were, however, also interesting subject-specific relationships.
The results of this study have theoretical as well as practical implications. Of theoretical interest are the gained insights in the way students' learning activities are regulated by internal and external sources which are more or less manipulable. Practical consequences have to do with the way students' learning activities are rewarded in exams and with the recognition and manipulation of factors that inhibit or promote constructive learning in our universities.

ESCAPING FROM MASS PRODUCED HIGHER EDUCATION

C.H.J. van Veen, H.K. Letiche, Z.C. Huijsman, & P. Jeffcutt,
The Rotterdam School of Management, Erasmus University,
Rotterdam, The Netherlands

Summary

The Rotterdam School of Management, Erasmus University, has been faced with an extreme growth rate in the number of first year undergraduates. In 1984 the total number of first year students was limited to one hundred, now in the academic year 1991/1992 seven hundred freshmen were allowed to register. This growth will in principle continue up till 1995 when a total number of one thousand freshmen are anticipated.

Taking into account the austerity budgets which Dutch universities are forced to work with, the consequences for the first year curriculum are enormous. Small scale education has become difficult to maintain and has increasingly been replaced by large scale lecture courses, which in fact have become the norm in freshman education.

The problem becomes even more serious, when we take into account the principles which the Rotterdam School of Management was used to work with. Principles which have been drawn from a descriptive approach to "the manager" (vs prescriptive). The committee responsible for curriculum design and development defines the end terms students have to meet to graduate from the Management School in the three categories: cognition, skills and attitudes. The attitude factor is seen, in comparison with other majors, to be specific to the school; we note further that it is evaluated very positively by the students. However, due to the growth of student numbers and the inevitable tendency towards large scale lecture courses, attitude development requiring intensive staff–student interaction has become problematic and exceptional.

The curriculum committee has been concerned that the reduction in courses appropriate to attitude development will cause a major problem for freshmen for whom the first year is supposed to fulfil an "orientation" function, whereby students explore the appropriateness for themselves of the programme offered, and the Management School tests their ability to meet the demands of the curriculum. Because of the ever increasing scale of the programme, the first year's curriculum no longer adequately reflects the need for attitude development, characteristic for a programme which emphasizes *development to become a manager.* It is feared that the students best adapted in their attitudes to become a manager will leave the programme, because of failed "orientation" phase, and in this way will be lost

1125
311

to the School of Management. Students who are seen as active, initiative takers and enthusiastic are favoured by the School.
With this background, the curriculum committee has defined an experimental project with the mission statement: "To offer students an orientation to academic work and attitude development, both being understood to be essential to management education". In order to challenge students who are genuinely interested in the field of management, the project has been defined as an extra curricular activity in which the student chooses to be involved in his/her orientation to the study. Student participation is voluntarily, without any sanctions to non participating students and/or rewards in terms of grades, credits, or what ever. Orientation and attitude development are made possible by offering students the chance to participate in work groups where they can discuss and study the field of management. Central to the project is the opportunity to work in small groups with an emphasis on individual initiative, responsibility and team work. Based on the objectives concerning attitude development, and offering an orientation based on the descriptive approach to management, the theoretical framework chosen for the project is an interpretation of Interactive Experiential Learning (Letiche 1986).
The essence of this interpretation of Interactive Experiential Learning is, that next to the abstraction and observation involved in cognitive activity, students have to interact contextually, prompting social learning. Experiential learning is not seen to be a prescribed technique, but is constructed to be made up of a dynamic tension between different elements, i.e. learning by doing, learning how to learn and learning through experiential group processes. The interaction between these different elements achieve a learning process which is rich in experiential value, and heuristically powerful enough to make enquiry successful. The paper will *in–depth examine* the interactions between the afore mentioned elements, and the way in which staff and third year students are able to act as facilitators to this process.
Besides providing a unique learning opportunity to the first year students, the experiment offers third year students an opportunity to act as facilitators and to learn to set up as well as participate in an Interactive Experiential Learning environment. A selected group of third year students have attended a course in management education, which prepared them to participate in the "orientation" project. During this course they were placed in a situation preparing them to facilitate experiential learning with first year students. Two members of staff have supervised the learning process of the third year students, and thereby have provided indirect guidance to the first year students.
The first phase of the project was started in September 1991 with the training of the third year students as facilitators. The second phase started in November 1991, when the third year students took over the supervision of more than two hundred freshmen (out of a population of seven hundred) who applied to participate in the project, resulting in a final presentation at the beginning of April 1992.
The evaluation of the experiment will be conducted under the responsibility of the curriculum committee, in cooperation with the members of staff. In essence the evaluation will determine whether the goals of the experiment (set by the curriculum committee) have been realized. The evaluation will assess the perception of first year students of the experiment to reveal whether they have been in contact with the specific defined management attitudes. Secondly has the

programme contributed to the "orientation" function of the first years, and specifically to the "orientation" of those students who are highly valued by the School of management, but might leave the programme because it stiffens initiative, and individual creativity.
A second evaluation will be conducted by the participating members of the staff in which they focus on the theoretical framework of the experiment, and evaluate if this framework has achieved what it was suppose to achieve. This evaluation will be based on documents by first and third year students produced during the process of the project, and notes on findings of the members of staff and observers who were invited during the project to observe its progress, and opinions of stakeholders within the School of Management. The outcomes of both evaluations will be used as inputs to discuss the difficulties faced and problems to be solved. The conference paper will evaluate the results of this project and will discuss possibilities, options and obstacles of *voluntary self–orientation* facilitated by older students with staff support.

References

Letiche, H.K., (1986). Interactive Experiential Learning in Enquiry Courses, paper presented at Developing Enquiry–Based Courses:Further Concerns, Newham College Cambridge: U.K.

INTERNATIONALIZATION OF EDUCATION: A FOREIGN LANGUAGE AS THE MEDIUM OF INSTRUCTION

Diana Vinke, & Wim Jochems, Faculty of Philosophy and Humanities, Delft University of Technology, Delft, The Netherlands

Introduction

University networks are being consolidated and expanded, both on a national and international level. Internationalization of education itself is becoming more and more prominent; it may lead to introducing a foreign language - for example English - as the medium of instruction in regular higher education, if only on a limited scale for the time being. This may raise the issue whether learning in a foreign language affects academic performance.

At first sight it seems plausible to assume that students need to have an adequate command of a foreign language if education is conducted in this particular language. Surveys in the United States and Great Britain reveal that international students as well as their teachers consider language and communication to be a major problem of studying abroad. Nevertheless there is still no generally acknowledged theory on what components (foreign) language proficiency consists of, nor on the precise nature of the relationship between foreign language proficiency and academic success. Theoretical frameworks in this area, such as presented below, are tentative by nature.

Foreign language proficiency and academic success

Canale (1984) proposed a framework which takes into account three dimensions of language proficiency: basic, communicative and autonomous. In addition, Canale identified four areas of knowledge and skills involved in each of these three dimensions: grammatical, sociolinguistic, discourse and strategic competence. In his view a widely used language test such as TOEFL (Test of English as a Foreign Language) would be inadequate for screening purposes because it mainly focuses on basic language proficiency, whereas international students' academic performance in an English-medium instructional setting relates to autonomous language proficiency.

Cummins (1984) developed a framework that relates language profiency to academic achievement among bilingual students. According to his framework language proficiency may be expected to vary along two continuums: context-embedded versus context-reduced and cognitively-undemanding versus cognitively-demanding. He concluded that, in general, the more context-reduced

and cognitively-demanding a particular language test, the more language proficiency (as measured by this test) will be related to academic success. This implies that international students' academic performance would be sensitive to their level of English proficiency, since classroom settings are mostly context-reduced and cognitively-demanding.
A review of empirical studies into the academic success of international students at English-medium universities (Jochems, 1991; Vinke & Jochems, 1992) shows that learning in a foreign language does affect academic performance. For those international students whose command of English is not adequate yet, academic load is reduced; courses of the regular programme are replaced by a corresponding number of language courses. On the whole the level of foreign language proficiency required for academic success tends to be lower for technology and natural sciences than for social sciences or the humanities. In addition, the lower the level of foreign language proficiency, the more important a factor it becomes in determining academic achievement. It is not clear yet what kind of language skills in particular international students need to have mastered in order to successfully complete studies at an English-medium university. Finally one should realize that - irrespective of the language problem - factors such as mathematical ability and student effort may affect academic success as well.

International students in postgraduate engineering education
This paper presents the empirical findings of a study into the academic achievement of a sample of 90 Indonesian engineers. These subjects took a one-year postgraduate course at the International Institute for Hydraulic and Environmental Engineering (IHE) in Delft. They were prepared for this English-medium course in Sanitary Engineering by an intensive eight-week English course, which they took in Indonesia and which was concluded by TOEFL. TOEFL scores were a measure of English proficiency and academic success was measured by averaging seven examination scores. In order to examine the relationship between English proficiency and academic success, Pearson product-moment correlation coefficients between TOEFL scores and exam score averages were estimated for the total sample and for TOEFL-subgroups.
For the sample as a whole this correlation coefficient is 0.51, which means that to a certain extent a lower level of English proficiency tends to go together with less academic success. Moreover, a t-test showed that engineers with TOEFL scores below 450 - the mean TOEFL score of the total sample - earned significantly lower exam score averages than engineers with TOEFL scores of 450 and above.
Further analysis revealed that, on the whole, completing the programme successfully and qualifying for a diploma required an additional effort from the low English proficiency group. Additional student effort was measured by the percentage of repeated exams. Engineers with TOEFL scores below 450 appeared to have repeated twice as many exams as engineers with TOEFL scores of 450 and above (18.7% and 9.1%, respectively). Moreover, all of the repeaters in the latter category qualified for a diploma in the end, compared to 74.3% of the repeaters in the former category. The overall success rate of the sample was 84.4%; this pass rate rose to 97.6% for the high English proficiency group and dropped to 73.5% for the low English proficiency group.

1129

The findings above are congruent with the results of some previous studies of this kind. However, a more remarkable observation is the relationship between English proficiency and academic performance within the TOEFL-subgroups. For the high proficiency group the correlation between TOEFL scores and exam score averages equals the correlation found for the total sample (r=0.51 in either case), whereas this correlation comes close to zero for the low proficiency group (r=0.12). Put differently, an improvement of English proficiency actually increases the chance of being academically successful only for students with TOEFL scores of about 450 and above.

The correlation between English proficiency and academic success might be explained as follows. There is a range of TOEFL scores within which a better command of English increases the chance of academic success to a certain extent and within which a limited lack of English proficiency can be offset by greater student effort or greater academic abilities, for example mathematical ability. This range has a lower as well as an upper limit. With TOEFL scores either below the lower limit or above the upper limit of this range, an improvement of English proficiency does not contribute substantially to a better academic performance. In the former case the lack of English proficiency is a real impediment to such an extent that even a greater student effort or greater academic abilities will not increase the chance of academic success. In the latter case the level of English proficiency is such that it will hardly stand in the way of academic success (cf. Vinke & Jochems, 1992).

Implications

The findings of this study have several implications. First, they confirm the need for a detailed and empirically validated theory of language proficiency, on the basis of which a more adequate language test than TOEFL may be developed for admission purposes. Second, additional research is recommended to determine what factors other than English proficiency affect academic performance, so that admission criteria might be set accordingly. We suggest that admission criteria in engineering education include various combinations of minimum English proficiency and mathematics scores. Finally, we formulated the hypothesis that there is a range of TOEFL scores within which a better command of English increases the chance of being academically successful. The lower and upper limits of this range may vary according to academic discipline or may even be institution specific. We recommend conducting institution-specific studies and using the established lower limit as a cutoff score for admission to English-medium higher education. In view of our findings a minimum TOEFL score of about 450 should be sufficient for international (graduate) students at IHE, which is fairly low compared to current cutoff scores at American and Canadian universities.

References

Canale, M. (1984). On some theoretical frameworks for language proficiency. In C. Rivera (Ed.), Language Proficiency and academic achievement (pp. 28-40). Clevedon, Avon: Multilingual Matters Ltd.

Cummins, J. (1983). Language proficiency and academic achievement. In J. Oller (Ed.), Issues in language testing research (pp. 108-129). Rowley, MA: Newbury House.

Jochems, W. (1991). Effects of learning and teaching in a foreign language. European Journal of Engineering Education, 16, 309-316.

Vinke, D., & Jochems, W. (1992). English proficiency and academic success in engineering education (Delft Studies in Communication and Education No. 1). Delft: University of Technology, Department of Communication and Education.

1131

ACADEMIC SUCCESS OF FOREIGN STUDENTS IN ENGINEERING EDUCATION

W. Jochems, W.J. de Jong, Y. Rijns, H.J. Smid, & A. Verweij, Faculty of Philosophy and Humanities, Faculty of Technical Mathematics and Informatics, Delft University of Technology, Delft, The Netherlands

Introduction

Every year about one thousand of foreign students seek admission to Delft University of Technology. Two to three hundred of these applicants are accepted, which equals approximately 8 to 10% of the total number of freshmen. Those foreign students who cannot submit a Dutch certificate of pre-university education have to take a language test to prove that their level of Dutch language proficiency is adequate for academic purposes. In addition, if students cannot produce a certificate which - in agreement with international treaties - qualifies them to study at a university of technology, they have to pass exams in mathematics and physics in order to be admitted.

In the summer of 1990 research into the academic success of foreign students in engineering education was started at D.U.T.. A review of empirical studies into the relationship between student characteristics and academic success in higher education for both native and foreign students was carried out. The results can be summarized as follows.

First, general theoretical models relating student characteristics and student performance are available (Haertel, Walberg & Weinstein, 1983). With respect to student characteristics two factors are considered to be very important, namely ability and study efforts. Empirical studies have corroborated the impact of ability and study efforts on academic success (Fraser, Walberg, Welch & Hattie, 1987). According to Fleming & Malone's synthesis (1983), mathematical ability appears to be the most important factor for academic success in science.

Second, studies investigating the relation between language proficiency, mathematical ability and academic success for foreign students in engineering education are very scarce. As a matter of fact, only three relevant studies have been found (Sharon, 1972; Ayers & Peters, 1977; Light, Xu & Mossop, 1987). According to these studies foreign students' language proficiency is an important factor for academic success in engineering education, but mathematical ability appears to be even more important. More detailed descriptions have been

published elsewhere (Jochems, Rijns, Smid & Verweij, 1991a and 1991b; Jochems, 1991).

Foreign students in engineering education

At present our research concentrates on foreign students who started their studies at the faculties of technical informatics, electrical engineering and mechanical engineering between 1984 and 1989. In this paper we restrict ourselves to students of technical informatics. Their examination scores as well as their scores on the Dutch language and maths entrance tests were compared to scores of a control group of Dutch students.

A first analysis showed a striking difference between academic achievement of the foreign and the Dutch students. Of the Dutch students 77% completed their first year successfully, compared to 41% of the foreign students. Furthermore the successful foreign students needed much more time to complete their first year programmes: it took them an average of 34 months, whereas their Dutch counterparts managed to complete the first year programme within an average of 21 months.

The correlation between Dutch language proficiency, measured by the language entrance test, and academic performance, measured by first year examination scores, proved to be about 0.40. Moreover a statistically significant difference was found between the mean language test scores of those students who completed their first year successfully and those who did not.

The correlation between scores on the maths entrance test and academic achievement was about 0.30. As far as mathematics is concerned, entrance test scores are not available for all foreign students because of the admission requirements mentioned above; in addition, we have doubts about the reliability of this test. At present an alternative, more adequate test for foreign students is being developed. If scores on the language test and maths test are combined, the correlation between these combined scores and first year examination scores is about 0.50. This suggests that for foreign students the combined scores provide better predictions of academic success than the single scores.

By analyzing and comparing the examination scores, the number of tries and the order in which students took the various tests of the first year curriculum, we tried to establish in what respects foreign students differ in their study behaviour, compared to Dutch students. The most striking differences are that a significant number of foreign students postpone the actual start of their studies considerably, that they are also more inclined to postpone test taking, and that they often try to pass tests in an order not intended by the curriculum. Furthermore their test scores are worse than those of their Dutch fellow students, they need more tries, and often go on trying where Dutch students have already given up.

Implications

Our tentative conclusion is that the admission procedures for foreign engineering students might be more effective. Results indicate that requirements with respect to relevant previous knowledge, especially in the field of mathematics, should be stricter. In addition, foreign students have great difficulties in making an effective start with their studies. Therefore measures to improve their study behaviour should be considered as well.

At present we are investigating the academic success of foreign students at the faculties of electrical and mechanical engineering at D.U.T. to see whether the preceding findings are confirmed.

References

Ayers, J.B. & Peters, R.M. (1977). Predictive validity of the test of English as a foreign language for Asian graduate students in engineering, chemistry, or mathematics. Educational and Psychological Measurement, 37, 461-463.

Fleming, M.L. & Malone, M.R. (1983). The relationship of student characteristics and student performance in science as viewed by meta-analysis research. Journal of Research in Science Teaching, 20, 481-495.

Fraser, B.J., Walberg, H.J., Welch, W.W. & Hattie, J.A. (1987). Syntheses of educational productivity research. International Journal of Educational Research, 11, 145-252.

Haertel, G.D., Walberg, H.J. & Weinstein, T. (1983). Psychological models of educational performance: a theoretical synthesis of constructs. Review of Educational Research, 53, 75-91.

Jochems, W. (1991). Effects of learning and teaching in a foreign language. European Journal of Engineering Education, 16, 309-316.

Jochems, W., Rijns, Y., Smid, H.J. & Verweij, A. (1991a). Onderwijs door middel van een vreemde taal. Tijdschrift voor Hoger Onderwijs, 9, 165-172.

Jochems, W., Rijns, Y., Smid, H.J. & Verweij, A. (1991b). Mathematics, language, and foreign students. Proceedings of the 6th European Seminar on Mathematics in Engineering Education, Budapest, 61-66.

Light, R., Xu, M. & Mossop, J. (1987). English proficiency and academic performance of international students. Tesol Quarterley, 21, 251-261.

Sharon, A.T. (1972). English proficiency, verbal aptitude, and foreign student success in American graduate schools. Educational and Psychological Measurement, 32, 425-431.

OLDER STUDENTS IN HIGHER EDUCATION: AN INVESTIGATION INTO THEIR WANTS AND ATTITUDES

Erika Timmer, & Ruud van der Veen, Katholieke Universiteit Nijmegen, Nijmegen, The Netherlands

By the improvement of health care older adults live longer. At the same time early retirement from work is more common. Housewifes get, in there fifties when children leave home, a second chance to study. Many of these still healthy, early retired adults are interested in higher education. During the last decade or two a growing number of European universities did start projects that offer educational opportunities that pretend to be more adapted to the special wants of older students. The research discussed in this paper focuses on educational wants and attitudes of older students.

The research is founded in theory on socialization in old age. According to Hurrelmann (1989) the individual person is seen as an active processor of the surrounding reality. Because of differential influences and demands during the life span, people differ in educational opportunities they get. The educational wants and attitudes of older adults are essentially seen as an reaction to earlier (deprivation of) educational opportunities and experiences. To highlight this process Erikson's (1982) theory is elaborated for 'young' older adults. Young older adults look for new possibilities for development after the stage of life, characterized by Erikson as that of 'generativity', i.e where they were devoted to the building of an existence, securing an income, raising children. Retirement form work and family duties constitutes on the one hand a challenge to compensate for one-sided development in the stage of generativity but on the other hand a danger of hedonism and as a consequence detoriating human capacities.

The research has been carried out in the context of a project, administered by the "Gerard Brom Foundation" (GBS) at the Nijmegen university (The Netherlands), that offers higher education for older adults. The project started in 1988. Essentially the GBS has a double aim, both facilitation of participation of older adults in the regular, main stream higher education and offering unique, exclusive courses for older adults. In fact till now the organization of exclusive courses for older adults dominates its activity. The research was carried out in 1991 by the "Institute for the Study of Adult Education" at the Nijmegen university, in close collaboration with GBS. As said, the central question in the research was an analysis of wants and attitudes of the participants and other adults that showed an interest in the course offer of the GBS in the first three years of its existence. On the base of theory, sketched above, and after four exploring group interviews with in total 20 (ex)participants in GBS courses, a questionnaire was developed and sent

to students and other older adults that showed an interest in the project. Well over 50 % (n=318) responded.
As could be expected, it turned out that the group was rather heterogeneous, i.e. the participants differ a lot with respect to the level of involvement. In the paper more precise data will be mentioned, e.g. with respects to academic subjects preferred by older adults (philosophy, history, arts, etc.) and personal motives of older adults (stressing personal development) for participation in higher education. The data turned out to be in line with outcomes of similar research done in other European countries.
The wants and attitudes expressed bij respondents are confronted with the theory on socialization. Hurrelmann stresses the active processing of reality. The data confirm this hypothesis. Older adults strive purposively to broaden their insights and competences:

- many of them, especially the women, confirm that their personality still changes and that their self-trust is growing through their study,
- most of them express a need to catch up and to refresh knowledge, independent of the amount of earlier education they had,
- most of the participants don't feel that they are unwanted in higher education, because higher education is directed to vocational education in the first instance,
- in contrast to the factual policy of most projects for higher education for older adults, respondents see no obstacle in working together with younger students; at the contrary respondents see benefits and chances in inter-generational dialogue,
- only some categories, especially the oldest women and students with low preparatory education, express a need for incidental provisions, with a didactical approach particularly developed for older students.

The discussion lead to the conclusion that the existing offer of the university of Nijmegen and of other universities in the Netherlands is rather limited in relation to the wants of older students. Dutch universities till now concentrate too much on unique, special courses for older adults, while many of the participants of these courses like to engage more actively in higher education. There seems to be "a market" for packages of courses, selected from the regular curricula in areas older adults prefer (philosophy, history, arts, etc.). The inter-generational character of this form of education seems, at least from the perspective of most older students, more likely an advantage than a disadvantage.

References

Erikson, E.H. (1982). The life cycle completed. New York/London: Norton.
Hurrelmann, K. (1989). Einführung in die Sozialisationstheorie - über den Zusammenhang von Sozialstruktur und Persönlichkeit. 2. Auflage. Weinheim: Beltz.

ADULT ACCESS TO HIGHER EDUCATION IN SCOTLAND: THE SCOTTISH WIDER ACCESS PROGRAMME

Pamela Munn, Scottish Council for Research in Education, Edinburgh, Scotland, United Kingdom

Introduction

Access to higher education has expanded in most countries during the twentieth century and in Europe increased access remains an important goal. In France, for example, a target of 80% of children educated to the Baccalaureate level by the year 2000 has been announced: in Britain the intention is to increase participation from 15% to 30% by the end of the century. European attention has been fixed mainly, but not exclusively, on the increased participation of school leavers. Thus the focus on the Baccalaureate in France, the qualification for university entrance, and the debate in the United Kingdom about the appropriateness of A levels or, in Scotland, Highers, in encouraging young people to stay on at school and gain entrance qualifications for higher education. However, there has also been interest in encouraging adult participation in higher education. Perhaps because of altruism, or perhaps because of the declining numbers of 16-18 year olds in the population, universities and other higher education providers have been admitting increasing numbers of adults to degree level courses. Who counts as an adult student varies from country to country. In the UK adult students are those aged 21 or over on age of entry to a higher education institution. This distinguishes them from younger students who enter from school at age 17-20, sometimes with a year's break between school and university or college. Statistics from the Department of Education and Science (1992) indicate that there are currently over one million students in higher education, of whom 372,000 are aged 25 or over. Looking at age of entry, it is noteworthy that in Scotland, 27% of all entrants to full-time higher education in 1989-90 were aged 21 or over; and this percentage is projected to increase by the end of the century (SOED, 1991). Of course, universities in the UK have traditionally provided for adult students through extra-mural studies, for example, and through their own access courses. However, a number of innovative programmes are being developed to attract even greater numbers of adults to higher education. One of these programmes is the focus of this paper, the Scottish Wider Access Programme (SWAP).

SWAP presents a new and challenging approach to adult participation in higher education in Scotland. It involves all Scotland's higher education institutions in collaboration with further education colleges and local authorities in providing guarantied access to higher education for adults (those aged 21 or over) who

1137

successfully complete an access course. The access courses are provided by further education colleges and students receive a grant to attend from their local authority. The SWAP courses are challenging in the curriculum and assessment methods they use to prepare students for entry to higher education. The courses are designed on the basis of a modular curriculum, whereby units of work of 40 hours are designed with clearly specified learning outcomes (about 22-24 modules comprise the course). Assessment is continuous and criterion-referenced. Yet these courses are preparing students for degree level study where abstract, theoretical knowledge is valued, assessment is norm-referenced and is carried out at the end of term or end of year. There are SWAP courses in most curriculum areas although science, mathematics and engineering dominate, reflecting perhaps the vocational thrust behind the initial SWAP development.

The research

The study is a small-scale qualitative investigation of staff and students' perceptions of the adequacy of SWAP courses as preparation for higher education. Students on access courses were interviewed using a semi-structured approach to investigate their views on modules, teaching and assessment methods, and on other aspects of their access course. They were followed from access to their first year of participation in higher education and interviewed again to discover their experience of higher education and they were asked to reflect on how well their course had prepared them for higher education. In addition standard demographic data were collected.

Staff on access courses and staff in higher education were also interviewed about the adequacy of the courses, about the kinds of collaboration between further and higher education which had taken place to develop the courses and about the general policy of encouraging greater adult participation in higher education.

The research sampled science, mathematics, technology and engineering access courses on the one hand, and social science courses on the other. Social science data are still being analysed. The results below concentrate on science access. A total of four science access courses were researched, and 48 science students interviewed, 14 science access staff were interviewed and 14 science higher education staff interviewed. All interviews were tape-recorded and later transcribed. The average length of interview was about one hour.

In addition to interview, access course documentation was collected together with relevant documentation from further and higher education institutions.

Results

In general, students are well satisfied with their preparation for higher education. They have some doubts, however, about a modular curriculum. These will be discussed in the context of previous research on the problems adult students face in science courses in higher education. These problems include the subject knowledge required, the ability to organise independent study and personal problems such as the lack of child-care provision and financial support. Higher education institutions' responsiveness to these problems will be described.

The results will also be presented in a broader policy framework. Previous work on access to higher education has suggested that when access expands it is 'socially controlled by institutional differentiation so that élite universities remain the cultural

possession of traditionally advantaged groups, (Halsey, 1992). It is important to ask, therefore, how access students are distributed across Scotland's higher education institutions. Furthermore, science, mathematics, technology and engineering degrees have traditionally been less popular with adult students than social science or humanities. These degrees have tended to be the preserve of young, male, full-time students. Can access students' experience challenge this preserve? Fundamentally, however, we are seeing in Scotland the emergence of two divergent higher education cultures, one, traditional and élite, valuing theoretical knowledge, a norm referenced approach to examinations and organised on the basis of full-time attendance at a university; the other valuing applied, practical knowledge, using continuous, criterion-referenced assessment provided in a variety of modes of attendance, such as full-time, part-time an open learning. Can these cultures be reconciled and will current higher education policy make the process more difficult? The abolition of the division between university and polytechnic pull in one direction, while the rank ordering of institutions in terms of performance criteria in research and teaching pull in another.

References

Halsey, A. H. (1991). An International Comparison of Access to Higher Education Oxford Studies in Comparative Education, 1.

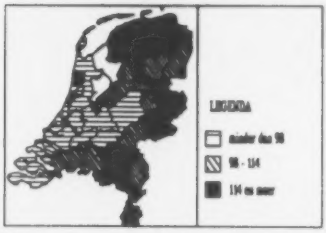
Department of Education and Science (1992). Education Statistics in the United Kingdom 1991 Edition Statistical Bulletin 3/92.

Scottish Office Education Department (1991). Scottish Higher Education Statistics Statistical Bulletin J2/1991/12.

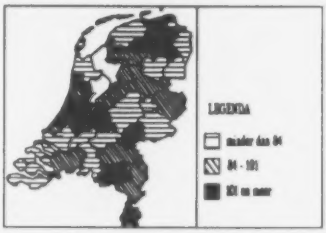
REGIONAL DIFFERENCES IN AGGREGATE ENROLMENT LEVELS IN HIGHER EDUCATION

Marc Vermeulen, TNO Institute for Spatial Organization, Delft, The Netherlands

Enrolment in higher education has increased immensely during the past 25 years or so. In 1960 61.8 per 1000 people aged 18–25 participated whereas this quantity was 161.5 per 1000 in 1986. Higher education is popular for many reasons: better chances on the labour market, higher income, attractive jobs etc.. Yet despite this strong increase, considerable differentiation in enrolment level subsists between regions.The explanation of these differences between regions in enrolment in higher education will be the central issue of this paper.



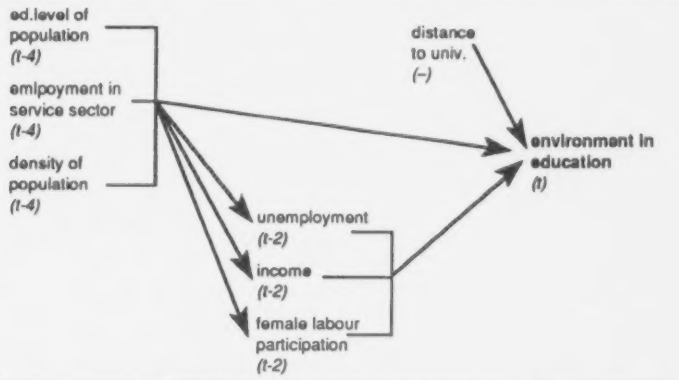
map 1 Regional differences in enrolment. Vocational Colleges, 1985/86 men & women.



map 2 Regional differences in enrolment. Universities, 1985/86 men & women.

From an individual point of view the choice to continue someone's educational career can be explained from factors like employment chances, expected and foregone earnings and costs of education. Apart from these more or less economic factors, educational choices obviously are influenced by socio–cultural factors such as gender, social class and cultural participation of the family. Extensive research has been done on these individual processes. Important theories in this respect have been formulated such as the human capital theory and the screening theory. In the public debate relations established on the individual level are often used to explain macro educational expansion as well. Yet empirical evidence on

1140



Scheme 1 A causual model to explain regional differences in enrolment in education

the aggregate level of analysis is weak. Relations that appeared to be important at the individual level turn out to be less significant at other levels of aggregation. As could be expected individual processes do not simply add up to macro relations. To explain educational expansion at the aggregate level one should conduct separate analysis within its own conceptual framework, with suitable data sets etc.. Research on the aggregate level is often done by the use of time series or by international comparison. We do not have an extensive tradition in this respect in the Netherlands. Historical and international comparison are often hampered by all kinds of definition problems and changes in systems.

Another way of analyzing aggregate enrolment in education is to examine regional differences in participation. As in historical and international research I use aggregate data on enrolment but I do not have the problem (at least in the Netherlands) of different system characteristics, definitions etc.. Analysis of regional differences also can be considered as 'some where between' the individual and the national level of analysis and might shed some new light on existing theories.

In this paper we will try to explain aggregate enrolment per region from regional characteristics such as economic conditions, the opportunity structure, urbanisation and location and socio–cultural features of region. (see scheme 1). The analyses will be carried out at the level of so–called COROP–regions (40) for the period 1976–1985. As a criterium to regionalize enrolment I use the region in which students lived at themoment they signed up for higher education. I will try to explain the total level of enrolment in higher education by means of a more ore less 'classical' regression model. In the model (scheme 1) I distinguish between three blocks of explanatory variables. At the most left hand side we have some background characteristics of the region. From a classical social status theory I use the proportion of higher educated people in the total population. Modern societies are characterized by a high level of service sector employment. In these white collar jobs level of education may be higher. I include labour volume in service

1141

sector as a percentage of the total labour volume per region. Finally I expect the density of population to have a positive effect on enrolment because of anticipated effect of urbanization. In the second step I introduce some specific labour market variables. Unemployment is supposed to have a positive impact on enrolment, because students can not find a job and try to get the better chances by continuing their education. Income may have a positive effect as well indicating that in rich regions people can afford to study (longer). Female labour participation will have a positive effect because it will be an indicator for some kind of emancipation. Apart from this we expect the effect to be higher in the equations for female enrolment. If female labour participation is accepted girls may anticipate by attending extra education. The final indicator I include is the average distance from the region to the nearest university. This variable should have a negative impact on enrolment. The presence of a university makes it easier to enrol in university.

Some problems will occur in the specification of so called time–lags in the model. Although we now what variable comes first and which is next from the causal model (scheme 1), I have to estimate the period between two stages in the model. I choose an empirical solution by fitting models with all possible (two year) time lags. The smallest time lags between the different stages appeared to be the most satisfying.

The model was tested on enrolment in 1985 as well as in 1982. It does explain enrolment in university quit well (adjusted R2 in 1982 0.53 for men and 0.70 for women and in 1985 0.64 for men and 0.70 for women). By far the most important variable is educational level of the total population: it has a strong positive impact on enrolment as well for men as for women. Population density some times has a small positive effect especially on female enrolment. Regional differences in service sectoremployment do not have any effect on enrolment in university, neither does unemployment have an effect. Income differences do not play an important role as well. Only for male enrolment in 1985 I found a small positive effect. Female labour participation has an effect I did not expect: the lower female labour participation, the higher enrolment in university in 1985. This effect goes as well for men as it does for women. In 1982 the effect is absent. Finally the presence of an university at a reasonable distance does not heighten enrolment in university.

I introduced one extra variable in the model to explain enrolment in vocational colleges (HBO), viz. the infrastructure of facilities in the expectation that regions without or with only a limited number of institutes will have low enrolment in vocational colleges. The model performs far worse on the explanations of enrolment in vocational colleges then it did on university education (adjusted R2 men 1985 0.30, women 0.33, 1982 men 0.25 and women 0.16). The proportion of higher educated people in the total population has in all cases strong direct positive effects on enrolment. These effects are toned down by indirect negative effects amongst others through the negative impact of income on enrolment. The negative impact of income has become less important in 1985 then it was in 1982. The opposite goes for the negative impact of female labour participation. It did not have significant (though negative) influence on enrolment as well for men as for women in 1982, but it did have a rather strong negative effect for both sexes in 1985. Finally I should point at a remarkable finding of the educational infrastructure: the number of vocational college facilities does not have any effect

on the enrolment level. The distance to a university however has a small but significant and positive effect on male participation: regions far off from university have higher enrolment levels for men then those regions with a university in their direct neighbourhood.

	Vocational Colleges				University education			
variable (jear)	1982 men	1982 women	1985 men	1985 women	1982 men	1982 women	1985 men	1985 women
distance to university (---)	0.44	0.00	0.33	0.00	0.00	0.00	0.00	0.00
HBO instit. (---)	0.00	0.00	0.00	0.00	—	—	—	—
income (t-2)	-0.49	-0.47	0.00	-0.42	0.00	0.00	0.20	0.00
unemploym. (t-2)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
female abour partic. (t-2)	0.00	0.00	-0.79	-0.62	0.00	0.00	-0.46	-0.30
educ.level population (t-4)	0.69	0.23	0.13	0.14	0.74	0.66	0.67	0.67
employment service sect. (t-4)	-0.43	0.12	0.00	0.00	0.00	0.00	0.00	
0.00 population density (t-4)	0.00	0.00	-0.26	-0.20	0.00	0.26	0.12	0.22
adjusted R^2	0.25	0.16	0.30	0.33	0.53	0.70	0.64	0.70

Table 1Regression results: Total effects.

1143

WHO STIRS UP THE INFLATION OF GRADUATION IN HIGHER EDUCATION?

P. Koopman, Amsterdam, The Netherlands

Introduction

In this paper we like to answer the question: "Which characteristics differ significantly for the several groups of students and graduates, who (intend to) continue higher education after graduation and groups who have no intentions for further higher education?"

In the eighties a trend towards permanent schooling can be seen in Dutch higher education. That is, graduation in vocational colleges (HBO) or university education (WO) is more and more an intermediate station.

We have the following three data sets.

Firstly, within the research group Studycareers Higher Education of the Foundation for Educational Research (SCO) of the University of Amsterdam we have at our disposal of studycareers of graduates in regular higher education (vocational colleges as well as university education) in 1987. In 1989 we have sent these ex-students a questionnaire. We asked, among others, about their future plans towards studies or courses on the level of higher education.

Secondly, we have at our disposal data on studycareers of students, who started a study at the Dutch Open University in 1988. In 1991 we have sent these students an extensive questionnaire. Among others we mapped their previous studycareers. Originally the Open University was mainly founded as an educational institute for second chance education (OU, 1990). But to a greater extent students at the Open University are previously graduated (in vocational college or university education). Between 1984 and 1989 their portion increases from to 51% (Ou, 1990, p. 37).

Thirdly, we have data on students who enroled the university in 1986. Of this group 10% is already graduated in vocational college (De Jong e.a., 1991). In 1989 we have sent all students an extensive questionnaire.

Thus we have data on three groups of graduates who continue higher education and their counterparts: stackers and three groups of non-stackers:

- graduates in vocational college who (intend to) continue higher education, and those who don't;
- graduates in higher education who started a course at the Open University, and students at the Open University who aren't graduated;
- starters in university education who are already graduated in vocational college, and those who aren't.

The comparison of these six groups will focus on the description of background characteristics (gender, age, parental educational level), (pre-college)school career, former secondary and higher education, amount of exact subjects in secondary education, delay), goal commitments and intentions.

Method
By the comparison we will mainly use comparative statistics, co-variance analysis.

Results
The analyses will take place in the spring of 1992. Anyway, we will join the papers presented at the ORD91 "The studycareers from male and female students at the Open University" (De studieloopbanen van mannelijk en vrouwelijke studenten aan de Open Universiteit) by U. de Jong and P. Koopman, and "Learning routes of students in university education" (Leerwegen van studenten in het wetenschappelijk onderwijs) by J. Roeleveld, U. de Jong and P. Koopman. Both articles are part of the ORD91-volume "Higher education and adult education" (Hoger onderwijs en Volwasseneneducatie, J.K. Koppen e.a. (eds), Amsterdam, SCO, 1991.)

Conclusions
At this moment the conclusions are unknown.

Discussion
The discussion will depend on the results.

References
Ou. (1990). Het onderwijs van de Open Universiteit: Ideeën, plannen en realisatie, 1980-1990. Zelfstudierapport ten behoeve van de eerste instellingsvisitatie, p.25. Heerlen: Auteur.

Jong, U., de, Koopman , P., & Roeleveld, J. (1991). Snelwegen en slingerpaden in en om het Hoger Onderwijs. Achtergrondstudies hoger onderwijs en wetenschappelijk onderzoek nr. 10. Den Haag: Ministerie van Onderwijs en Wetenschappen.

1145

WHAT STUDENTS EXPECT TO DO AFTER 16: THE SCHOOL'S ROLE

Ray Sumner, The National Foundation for Educational Research in England and Wales, United Kingdom

Background

All secondary schools in the UK have participated through local education authority (LEA) based projects in a national programme, the Technical and Vocational Education Initiative (TVEI), funded by central government's Department of Employment. The initiative aims to promote for students aged 14 to 18 years continuity and progression; which pays regard to breadth and balance in the curriculum, acquisition of learning skills including IT, preparation for the world of work, equality of opportunity, modern technologies, and personal development. A key objective is to increase the numbers continuing voluntarily in education or training beyond the age of 16, and in the longer term beyond the age of 18 into higher education or advanced training. The study outlined here took place in a large LEA which managed the programme's additional resources via First Phase school and college consortia (5 districts starting in 1988) and similar Second Phase consortia (8 districts starting in 1989). Students approaching the end of compulsory education at age 16 are Year 11; those enroled subsequently are in Year 12 and Year 13. There were 3 surveys; of First Phase Year 11 in 1990; a follow-up of First Phase Year 11 students who became Year 12 in 1991; and of Second Phase Year 11 in 1991. Additionally, small groups of the Year 11 sample were interviewed halfway through Year 12.

The First Phase Year 11 Survey

The principal aim of the study was to examine students' views on the guidance and experiences they had considered in taking `the next step'. About a twenty percent sample of students (N = almost 1200) responded to a questionnaire. Administration took place after all schools had completed trial external examinations in May 1990. Each student was asked on the basis of their current intentions to respond to either a 'stayers' section or a 'leavers' section. In the event, almost a tenth of the sample completed both sections.

All of the students gave their General Certificate of Secondary Education entries. Those intending to stay answered questions under the headings of 'What I hope to achieve' (15 items), `What I expect to do' (18 items), `Aspects that have helped me so far' (21 items) and `The course I hope to take next year' (an open-ended item). Over 750 students chose to answer this section of the questionnaire. The

remainder (about 360) answered the last section (15 items) concerned with reasons for not wishing to continue in full-time education. The findings are outlined below.

GCSE Entries

A substantial gender bias along the customary lines was found among the optional subjects. More males were taking Design Technology/Communication, Electronics, Chemistry, Physics, Information Technology, and Computer Studies; more females were taking French, German, English Literature, Home Economics, Typing, Business and Information Studies, Combined Science, and Religious Education. On the whole, the students intending to continue in education were entered for more subjects than those proposing to leave.

Expected Post-16 Achievements

Two-thirds of the 'stayers' said they wished to gain qualifications to enter employment; these included many of the 44 percent who wished to qualify for higher education admission. Only 15 percent of the male and 29 percent of the female students wanted to get a particular vocational qualification.

More female (48 percent) than male (32 percent) students gave high priority to developing skills in working with others, whilst more males (26 percent) than females (13 percent) gave top priority to acquiring computing or electronics expertise. About 20 percent of the group gave high priority to developing design capability. More females (13 percent) than males (5 percent) gave high priority to developing foreign language speaking skills. Only a fifth gave high priority to controlling their own progress, and lower proportions felt they should have a measure of balance in their course mix, or take part in clubs, societies, or community service.

Activities in the next stage

A third of the students thought they would attend a Further Education College (mainly vocational). Only minorities thought they would have laboratory (males 32 percent, females 25 per cent) or workshop (males 22 percent, females 11 percent) experience. Less than a fifth thought they would have work experience or be involved in events with employers. Only a tenth anticipated contributing to a college or school production.

Gender differences were evident in respect of playing in sports teams (males 33 percent, females 18 percent), learning electronics (males 14 percent, females 2 percent) and using drama facilities (males 5 percent, females 12 percent). There were marked differences, too, with regard to reviewing progress with a tutor (males 25 percent, females 47 percent), keeping one's own progress records (males 21 percent, females 36 percent), and completing a careers guidance questionnaire (males 20 percent, females 27 per cent).

Aspects facilitating continuity

About two-thirds of the group said that work experience had been most helpful and in the same vein, a similar proportion identified finding particular subjects interesting. Other aspects were: trial examination results (45 percent), parental advice (45 percent), college brochures (males 40 percent, females 60 percent),

college open days (males 33 percent, females 50 percent), careers officers interviews (males 37 percent, females 48 percent), visiting a firm (about 29 percent), having course objectives (about 28 percent), and problem solving (about 22 percent). Relatively few students found profiling or discussing their Record of Achievement to have been helpful or careers lessons/advice; however, about 25 percent thought that guidance from a tutor had been influential.
More female students felt they had been helped than males. Two-thirds of the helpful processes were controlled directly by the schools or colleges (e.g. having course objectives, careers lessons, tutorial guidance, open days, link courses, work experience). The remainder were in the students' ambit (e.g. interest in a subject, own progress criteria, parental advice, discussing with friends).

Next year's course
Amongst the 370 males and 390 females expecting to continue, the main options were: Advanced Levels (academic) 54 per cent; Business or Technician course 21 per cent; occupational courses 15 per cent; 10 per cent no reply.

'Leavers' intentions
There were 221 males and 137 females intending to find work or a training course. Sharp gender differences were evident, with 50 percent of preferences given by females following accustomed lines (e. g. shop assistant 14 percent, office work 18 percent, hairdressing 12 percent). Only about 8 percent of females saw the Youth Training Scheme as a viable next step compared with 16 percent of the males. Most of the males intended to find work which included training (e.g. vehicle mechanic 12 percent, electrician apprentice 8 percent, construction industry 11 percent). Twelve percent of the males and 17 percent of the females expressed no kind of intention except for one who said `dole'.

Reasons for not continuing
The most prevalent reason was `Experience is the best way of training' (males 70 percent, females 61 percent); apprenticeship combining training with earning (males 60 percent, females 44 percent) and the impact of work experience (about 59 percent) were also influential. Sizeable proportions wanted to be treated as adults (about 44 percent) and said they needed a wage (about 41 percent). Thirty-six percent said they already had a job to go to.
Significantly, 44 percent of the males and 19 percent of the females thought they would not work if they were to enrol at a college. About 11 percent thought the courses would not be interesting or that poor qualifications would debar them from entry. Others among the leavers (males 12 percent, females 19 percent) thought that college courses could prepare them adequately for starting a job.
The Youth Training Scheme was regarded as providing good skills (males 17 percent, females 26 percent) and leading to a job (males 12 percent, females 17 percent). The training allowance was regarded as extremely important (males 16 percent, females 10 percent).

The Group Interviews

To ascertain whether or not students' particular school experiences supported or disproved the survey findings, five groups of students were interviewed at their post 16 institution. For example, in the 4th and 5th years, were there careers lessons? which teacher took them? what other careers opportunities did they have? did they complete a careers questionnaire? and so on.

These discussions substantially confirmed the survey findings. For example, careers education seemed to have been very uneven in that some students could describe well co-ordinated activities shared between careers teachers, tutors and careers officers, whilst others merely viewed careers guidance as disorganised and inconsequential. The good quality careers education was highly regarded, though some of the Advanced Level academic students said that it had been of little help in considering possibilities in higher education and beyond.

Several students had changed courses during their first term in Year 12. These thought that better information about what would be required of them when studying for subjects or courses would have been of most use.

Gender bias or equal opportunities in course choice or occupation were not regarded as issues, as they had, apparently, never been raised with the students or their parents.

Records of Achievement had been introduced in the schools when the students were in Year 10. For many students, the experience had been off-putting, with little tutorial discussion, hurried and prescriptive personal accounts and a product with low esteem.

The non-examined courses provided by the colleges or schools were not seen as complementary to students' specialised or narrow lines of study (some students thought that A levels in two foreign languages and Mathematics were balanced and broad).

The Year 12 follow-up of Year 11 students

Three hundred and forty five students from the Year 11 sample were traced into Year 12. About a fifth of the original sample had left their course before completing Year 12. Almost 80 per cent were following Advanced Level (academic) courses, whilst the remainder studied technical or business courses. The majority were motivated by having chosen subjects or courses that interested them (mainly from pre 16 years). Other factors were: the future value of qualifications; good relationships with staff; teachers' enthusiasm; peers working hard; having modern apparatus; and lectures being well-prepared. Vocational course students emphasised regular assessment and feedback; progress records; tutorial reviews; and topics related to jobs. Transition had been facilitated by induction guidance procedures. Studies were enhanced by: having objectives for each stage of the course; progress reviews with teachers; individual use of learning resources; home study facilities; discussing problems with peers; and time at home to think things out. On the negative side, there was reduced interest in modern technology and IT, in learning a modern foreign language, and following a course with breadth and balance.

The Phase Two Year 11 Survey

Almost 1800 students from eight districts participated. Compared with the Phase One sample, more intended to continue into Year 12, with a significantly higher proportion opting for business or technician courses. Also, a much higher proportion said that Records of Achievement processes had been helpful. In most respects, however, the findings replicated the previous Year 11 survey.

Implications

- Tutor sessions provide schools and colleges with a range of opportunities for developing students' views of themselves in relation to present circumstances and future prospects. A vital task is to monitor each student's experience (e.g. students' check-lists for careers interview, syllabus objectives, tutorial guidance, etc.). One objective should be to produce action plans with each student (e.g. find out about the modules in a technician course).
- School and college policy on certain issues needs to be defined and declared (e.g. to raise gender bias with pupils, parents, employers, with a view to changing curricular options and work experience placements).
- Students' perceptions should form part of institutional review and development (e.g. Why do so many not consider it important to acquire a foreign language skill? design capability? IT skills? balance of arts, science, technology?).
- Guidance for students with unclear intentions should continue up to the end of Year 11, to encourage progression and continuity into Year 12 or an apt training course.
- The data illustrate the long-standing situation in England and Wales whereby discontinuity across much of the curriculum takes place in order that students may study a few subjects in some depth in order to qualify for entry to a higher education institution. To counter this constraint, students should be engaged in cross-curricular, vocational and personal development activities.
- The minority of students lacking confidence or regard for independent study using learning resources should be identified, encouraged and trained. Institutions would need to provide facilities for students who otherwise would not have access to them.

Acknowledgements

The surveys were organised and administered by the Director and staff of the TVEI programme in the County of Hampshire, with the collaboration of teachers and students in the institutions of the consortia in each of the phases. Data analysis was done by NFER Computing and Statistical services.

CAL ON DIGITAL DESIGN: STUDENT CONTROLLED PRACTISING & TESTING

Gerdy C. ten Bruggencate, & Jan T. van der Veen, Educational Centre; Martin F. Beusekamp, & Tom R.C. Bonnema, Faculty of Computer Science, University of Twente, Enschede, The Netherlands

Abstract

Computer Assisted Learning (CAL) has been used for more than ten years in a course on Digital Design for first-year students at the University of Twente. Recently a completely new version consisting of 27 types of problems has been developed. The new CAL program runs in the Microsoft Windows environment. The authoring system Inigo was used to produce the CAL materials. In the CAL program practising and testing are under student control. In this presentation a description of the CAL program and a report on the results of an evaluation will be given.

Introduction

For already more than ten years, Computer Assisted Learning (CAL) materials have been used in a course on Digital Design for students in the faculties of Electrical Engineering and Computer Science of the University of Twente. Originally, apart from lectures to present information, the course on Digital Design consisted of classroom sessions and a lab for practising. The classroom sessions were replaced by CAL exercises and tests to provide for more individualized and efficient learning.

The CAL system turned out to be satisfactory for both students and teachers. As the course on Digital Design needed to be modernized, a revision of the course, including the CAL materials, was carried out between January 1990 and August 1991. The new course was based on the book "Digital Design Fundamentals" by K.J. Breeding (Prentice-Hall, 1989).

In the CAL program on Digital Design several types of learner control have been realized. The use of computers in instruction allows for learner control where it would be impossible in a teacher controlled classroom. Learner control allows the student to control some instructional features, such as the sequence of topics in a lesson, the number of practice examples, or the type of feedback (e.g. immediate or called upon by the learner). In a literature review on "Learner Control in Computer-Based Instruction", Friend and Cole (1990) conclude that the effects of

many different types of learner control in CAL need more investigation. Learner control is supposed to be one of the elements which contribute to the effectiveness of a CAL program.The effects of learner control seem to differ according to the lesson content, the specific nature of the control options, individual traits of learners, etc..
In an evaluation of the CAL program on Digital Design, we investigated in which ways and to what extent learner control options were used by students. Our hypothesis is, that the learning process of students following the course on Digital Design will be optimized, when these students are enabled to control their own learning activities concerning practising and testing. However, this effect will only be achieved when students are sufficiently capable of managing their learning activities.

Description of the CAL program
Microsoft Windows was chosen as an environment for the CAL program, because of the multitasking facilities, the standardisation of Windows applications, its user friendliness and the possibilities of switching to the OS/2 environment in the near future. The CAL program was produced using the object-oriented authoring system Inigo. Inigo facilitates the development of CAL programs which are student controlled. The authoring system was developed by the Educational Centre of the University of Twente.
The CAL program consists of 27 types of problems, covering most of the basic skills and knowledge which are taught in the course. There are two modes of operation, the "practise" mode and the "test" mode. In practise mode, the program provides adequate feedback about the student's responses, to guide the student in solving the problem. In test mode, feedback is minimized. Students following the course on Digital Design are expected to practise each type of problem, until they judge themselves to be ready for a test. Solving each type of problem in test mode once is compulsory. Student control is incorporated in the CAL program in the following ways:

- The sequence of the problems can be determined by the student.
- Each type of problem can be practised (in different versions) as often as the student wants.
- While practising, the student is allowed to enter answers, until the answer is correct, or (in complex problems only) to ask the program to supply the answer.
- The student can run a problem in test mode when and as often as (s)he likes.

In addition to the types of control which were embedded in the CAL program, the student was also allowed to decide when and how often (s)he used the CAL program.
In the CAL program on Digital Design the student chooses the type of problem (s)he wants to solve from a menu. The program starts in practise mode. The student can select the test mode in a menu and use the same menu again to go back to the practise mode at any moment. As some input parameters for each type of problem are generated (quasi-) randomly, a large amount of versions of each problem can be produced. In practise mode, a "Generate button" is visible on the screen, to enable the student to start a new version of the problem. When an answer can be requested from the program, an "Answer button" appears on the screen.

In test mode, the results of a student are logged in a student file. Students can request an overview of their results via the menu. The student file is encrypted and it is protected by a password.
When the student has completed all the tests successfully, (s)he receives a certificate which is required to have access to the examination. An ample amount of computer time is reserved for each student during the course, but students are allowed to enter the computer lab and use the CAL material whenever they want during opening hours (day-time and evening) if there are computers available. On reserved hours students have priority to use a computer.

Method of evaluation of the CAL program
The evaluation of the CAL program was carried out between September 1991 and February 1992. To be able to evaluate the CAL material, not only test scores, but also the number of attempts on each test and the date on which a test was completed successfully, were registered in student files. Furthermore, student questionnaires were used to obtain information on appreciation of the CAL materials, user friendliness of the system, time spent on CAL sessions, the organisation of the course, etc.. In constructing the questionnaire special attention was paid to questions concerning student control.
Data were gathered of 213 first-year students in Electrical Engineering and Computer Science.

Results & Conclusions
Students were free to use the CAL program during a period of 8 weeks (the duration of the course on Digital Design). The logged dates show that there was a wide variety of starting dates throughout this period. In the questionnaire, only 20% of the students indicate that they started to work with the program at the end of the course. Most students finished the last test within two weeks after having started. According to student responses on the questionnaire, the time needed for practising and testing all types of problems was ± 7 hours. The average session time was ± 2 hours. It appears that most students started working with the CAL program in time and finished within a reasonable time.
The order in which the different types of problems were selected by students, was checked by looking at the dates of test completion. The students selected problems in the sequence that was presented in the menu, with only a few jumps back and forward. Student control of the sequence of problems appears not to be very important in this CAL program.
Most students needed no more than 1 - 2 test attempts on each type of problem, one relatively complex problem excepted. We can conclude that, although there was the possibility of a "trial and error" approach, this approach was not chosen by most students. According to the results on the questionnaire, 66% of the students practised solving each type of problem at least once, before starting a test. From the remaining 34%, a large group indicated that the practise mode was skipped only, when the new type of problem showed some similarity to a preceding type. In general, it seems that the extra feedback given in practise mode stimulates to adhere to the sequence of practising before testing.

1153

According to 87% of the students, the user interface (menus, buttons, way of answering, etc.) was easy to handle. Students did not have difficulties in controlling the dialogue with the program, probably due to the graphical interface. 81% of the students classified the CAL program as useful. Only 10% of the students indicated that working with the program took much time. Of all students, 79 % would like more CAL problems (on some topics which were still not covered by the CAL program).

Discussion

We conclude that the upgrading of the CAL program on Digital Design to the Windows environment has been successful. Students appreciate the CAL program as a useful and efficient way of learning. The control given to the student on the decision whether to practise or test a certain type of problem leads to a desired learning behaviour. Most students practise solving a problem (if necessary) before entering the test mode. We did not investigate how many versions of each type of problem were practised, before students entered the test mode, or how many times students requested an answer from the program. But as students usually needed no more than 1-2 test attempts, we can conclude that the practising was sufficient. It seems that under the condition of a compulsory test, students are capable of managing their learning behaviour adequately.

References

Friend, C.L., and Cole, C.L. (1990, November). Learner Control in Computer-Based Instruction: A Current Literature Review. Educational Technology, 47-49.

COMPUTER ANXIETY AT THE BEGINNING OF THE STUDIES

Päivi Kolehmainen, University of Joensuu, Joensuu, Finland

Introduction

The continuing development of technology has nowadays strong effects on schools and every kind of educational acts. People think mainly that computers are good teaching tools, but there are also many negative aspects related to the use of computers. These negative interpretations, like anxiety, occur usually in situations where one is learning something new and this causes the resistance of change. Computer phobia is defined as a) the resistance to talk about computers or even thinks about them, b) the fear or the anxiety towards computers and c) angry or aggressive thoughts towards computers (Jay 1981). In this study I, however, talk about computer anxiety instead of computer phobia, because a computer differs from other objects of fear. In computer anxiety the fear has a computer as an object. The fear is not only the fear towards a machine, but it has many unclear and mystic elements before proper knowledge and experience. It is important to reduce anxiety, because it has negative effects on learning.

Theoretical framework

Computer technology is getting more and more general nowadays. Thus, it is important to notice people's opinions about using computers. Do they think computers are controllable or not (See Weizenbaum 1984)? Human-computer -interaction is a complex wholeness that causes different kind of emotional interpretations. These interpretations depend on many factors like user's experience and activity, social support, the sense of control, coping-strategies, personal characteristics and the common effects of all these. Why is it then so difficult to get used to computers? People have to adapt to computer's ways of functioning instead of one's own thinking and because people do not think like computers, it causes trouble. Nowadays markets are, however, competing with user friendlier systems. Human-computer -interaction has been considered mainly from two different point of views, which are cognitive and emotional. In cognitive point of view we are interested in differences between human being's and computer's thinking processes (See Danzer-Kahan, Schwatlo & Dirlich 1986). Emotional point of view handles people's fears of unemployment caused by development of computer technology. Another fear is that computer technology makes working tasks poorer. Social support is also an important aspect in human-

computer -interaction, because it often makes the situation less anxious (See Briefs 1982; House 1981).
In human-computer -interaction one of negative interpretations is computer anxiety, which interferes the communicational nature of the human-computer -interaction. Computer anxiety can be seen as a specific form of test anxiety. Both forms of anxiety are usually related to performance situations and a person is calm outside the stressful situation. Test anxiety is defined as an unpleasant feeling or emotional state, which has psychological effects on behaviour and which occurs usually in testing and estimating situations (See Dusek 1980). A person processes too much irrelevant information that interferes concentration. One gets anxious especially in difficult performances, not so much in routine tasks. Computer anxiety is defined as a fear or a prejudice, which appears when one is using computer technology or when he/she is thinking the consequences of the use. A person gets anxious of the use, even though the computer itself is not dangerous at all (See Maurer 1983; Marcoulides 1989). It is also important to reduce computer anxiety, because it is harmful for learning and effective use of computers. To reduce computer anxiety, reasons for the anxiety (like equipment, software or environment) have to be clarified first. In that way the right kind of education can be planned. Many studies have indicated that an experience reduces computer anxiety. Because computer anxiety is not a stable feature but a form of state anxiety, it is our hypotheses that computer anxiety can easily be changed.

Method

The subjects of the study were 29 teacher students who participated the basic course in computer science in spring 1989. The course consisted of lectures and practices. In lectures there were presented subjects as computers in education, the position of computer technology in the future and new directions like artificial intelligence. In the end of the course the students had an exam. In practices the students first got familiar to the structure of computer equipment and an operating system. After that they used computers as a tool; they used word processing, spreadsheet and card index software.
The students answered a questionnaire consisting of three parts. The first part handled anxiety related to computers, the second attitudes towards computers and the third thoughts and feelings related to computers. It was used a pretest-posttest-arrangement in the study; the students answered the questionnaire both at the beginning and at the end of the course. This kind of pre-experimental arrangement has many methodological problems. First, is it really the course that reduces anxiety or could it be time effect or maturation between the pretest and posttest? The students may have learned something about computers also from elsewhere or they may have learned something from the pretest itself. There would be less problems if the arrangement would contain different control groups. There should be group that does not have the course between the tests and the groups that have only pretest or posttest. The data was analysed using factor analysis and one-way variance analysis.

Results and discussion

The main result of this study indicated that the basic course in computer science reduced computer anxiety. At the end of this course the students' own estimation of

their own knowledge was much higher than at the beginning of the course. Thus, the computer anxiety is reduced because of increased experience. The result is supported by other studies, which all indicated that the experience about computers reduces computer anxiety (See Asikainen 1990, Cambre & Look 1985, Kolehmainen 1991).

In this study the basic course in computer science reduced especially anxiety of two indicators: computer equipment and the use of computers. The course had also positive effects to attitudes towards new technology, especially to human-computer -interaction and to computers' position in the future. Negative feelings related to computers as the fear of failure and the desire to escape were reduced during the course. So, the course had positive effect also to the thoughts and feelings related to computers.

Computer anxiety and the attitudes towards computers can be studied also in many different ways than in this study has been presented. One possible way is to increase some qualitative aspects to the study and a person could tell in a deeper way how he/ she feels. How can the anxiety then be reduced? The people's experience should be increased in any way and all the borders in front of this experience should be eliminated. This is important, because anxiety reduces effective use of computers and has negative effects on learning.

References

Asikainen, E. (1990). Lasten ja nuorten suhtautuminen tietokoneeseen ja teknistyvään tulevaisuuteen. Kasvatustieteiden tutkimuslaitoksen julkaisusarja A 36. Jyväskylä.

Briefs, U. (1982). The Impacts of computerization on the working class and the role of trade unions. In Bjorn-Andersen, N. 1982. The Human side of information processing. Proceedings of the Copenhagen Conference on Computer Impact -78 October 25-27, 1978. Amsterdam.

Cambre, M. A. & Look, D. L. (1985). Computer anxiety: definition, measurements and correlates. Journal of Educational Computing Research ,1, (1), 37-54.

Danzer-Kahan, U., Schwatlo, U. & Dirlich, G. (1986). Die Gestaltung der Mensch-Computer -Interaktion Einfhrung in ein Interdisziplinäres Forschungsgebite. In Dirlich, G., Freksa, C., Schwatlo, U. & Wimmer, K. (Eds.). 1986. Kognitive Aspekte der Mensch-Computer -Interaktion. Workshop, München, April 1984. Springer-Verlag Berlin Heidelberg.

Dusek, J. B., DeYaeger, K. M. & Mergler, N. L. (1975). Brief Reports: Information Processing in Low- and High-Test-Anxious Children as a Function of Grade Level and Verbal Labeling. Developmental Psychology ,11, (5), 651-652.

House, J.S. (1981). Work stress and social support. Addison-Wesley, Reading.

Kolehmainen, P. (1991). Computer anxiety at the beginning of the studies. University of Joensuu. Joensuu (In Finnish).

Marcoulides, G. A. (1989). Measuring computer anxiety: the Computer Anxiety Scale. Educational and Psychological Measurement ,49, (3), 733-740.

Weizenbaum, J. (1984). Computer power and human reason. From judgement to calculation. Bungay, Suffolk.

THE TRANSFER EFFECT OF TEACHING SYSTEMATIC PROBLEM SOLVING TECHNIQUES IN COMPUTER ASSISTED LEARNING

E.L. Free, C.L.M. Pouw, & C. Terlouw, Educational Centre, University of Twente, Enschede, The Netherlands

Students in physical engineering are faced with many complex problems of various kinds. Amongst others, Electricity & Magnetism is one of the subjects in which a systematic problem solving technique attributes to a better problem solving behaviour and to a better understanding of the physical laws concerned. The problem addressed in this paper is which instructional functions have to be realized in a computer assisted learning environment for teaching problem solving. In a practical learning environment first-year university students had to solve a problem concerning a capacitor using a computer program. Just before this task the experimental group had followed a computer guided problem solving instruction regarding a similar problem, whereas the control group had no previous instruction.
The behaviour of both groups of students is discussed and compared with an expert behaviour on the same problem.

Objectives
Earlier research showed that the process of solving mechanics problems can be supported by the computer (Pouw & Staal, Staal & Ten Bruggencate). These programs have been designed on the basis of a system of heuristics through which the student is guided by the computer. As the main disadvantage of the computer support it was mentioned by the teachers that the students are too less trained in putting the right questions; the students are too much guided in this respect. Although the appreciation of the programs by the students was high, there appeared also to be the wish for more student controlled programs. The objective of the research reported here is to improve the methodology of the design of computer support in the learning of problem solving by students in higher education taking into account the above mentioned points of computer support and student control.

Theoretical framework
Our theoretical framework originates from the socialhistorical school, in particular the activity theory applied to instruction and instructional development (Vygotsky, Leont'ev, Van Parreren). The kernel of the approach is to be found in designing such educational situations as will enable students to acquire efficiently more

possibilities of orientation in new problem situations (Van Parreren). A larger orientation space will better enable students to prepare for successful execution of the problem solving process. Van Parreren distinguished some tasks in developing such instruction. Mettes et al. and Terlouw proposed a systematic procedure for instructional development in which the development tasks of Van Parreren are formulated in detailed heuristics. For our investigation the "design phase" is of particular interest, especially the "mini theory", prescribing how to teach problem solving in a specified domain to a specific group of students in a specified teaching-learning situation. The following main instructional functions derived from the mini theory are relevant for learning to solve problems: orientation, practising, testing and giving feedback. These functions will form the basis for the design of a computer supported systematic problem solving instruction in the field of Electricity & Magnetism.

Research questions

The following research questions are relevant:

- Which instructional functions have to be realized in what way in the design of computer support with a low or a strong student control for teaching problem solving in the field of E & M?
- Facilitates a previous computer controlled orientation on the system of heuristics for problem solving the solving of a student controlled problem?
- To what degree is student control an efficient way of practising computer supported problem solving?

Methods

Design of the computer support.

In order to enhance a systematic problem solving behaviour a computer assisted learning program has been developed in the field of E & M about the capacity of a capacitor. The students are faced with two in essence similar problems about the capacity of a capacitor partly filled up with a medium. But there is an essential difference in the amount of student control in the problem solving approach. In one of the problems the student has to follow a prescribed route through the program. This route is based upon a subject matter analysis in combination with heuristic schemes for problem solving in related fields (Van Weeren et al.). This exercise fulfils the instructional function of the orientation to a systematic problem solving approach in the specific domain of Electricity & Magnetism. The student has to exercise the domain specific problem solving process and receives an adequate feedback by the computer on his actions. The student is forced to follow a prescribed path in order to make the essential elements of knowledge and the actions, which are specific for this subject, operational.

In the other problem a more realistic situation is created in which the computer offers a learning environment with a variety of possible entrances towards the solution of the problem. The student is free to choose a problem solving approach that makes sense to him. The computer provides the student with declarative and/or procedural feedback only on request by the student or in situations in which it would otherwise become impossible for the student to solve the problem. This exercise fulfils the practising function in a student controlled environment. The

feedback which is given during the exercise depends on what outcome the student has reached and whether it satisfies a standard solution.

Experimental set-up

The second question addressed is whether a previous examining of the more computer controlled problem (orientation) facilitates the solving of the more student controlled problem (practising). In the educational practice of the curriculum of physical engineering all first-year physics students (n=100) have to follow the computer assisted instruction course. The experimental group (n=50) first passes the orientation in the computer controlled problem, after which the student controlled problem has to be solved. The control group (n=50) will start practising immediately with the more student controlled problem. The interactions of the students of experimental and control group with the program will be gathered in log-files, especially for the student controlled problem, in order to get insight in the entrances that students chose and the amount of help that is asked or needed by the students to solve the problem.

Efficiency of student control.

In order to obtain a reference for analyzing the collected routing data from the student controlled problem, some experts (teachers of electromagnetism) will be asked to solve the problem in the open learning environment while speaking aloud. Their interactions will also be recorded in log-files. Comparison of the routes chosen by the experts with those of the students in the experimental and the control group will give an insight in the efficiency of the student controlled program.

Results

The experiment will be conducted in the early spring of 1992. The first version of the computer program has been tested during the regular course of electromagnetism in 1991. A formative evaluation showed a great satisfaction in using the program. It was appreciated as a contribution to learning problem solving. A majority of the students preferred the student controlled problem above the computer controlled one.

The present research concentrates on the effectiveness of the computer controlled orientation function and the student controlled practising function.

PROBLEM SOLVING AND SEQUENCING FOR PROBLEM SOLVING IN MECHANICS IN HIGHER EDUCATION

J.W. Mulder, W.J.M. Tiemessen, & H.K. Slettenhaar, Institute for Higher Education, Enschede, The Netherlands

Introduction

The research presented in this paper concerned the improvement of the learning-outcome of a course in Mechanics, 1th grade, Higher Education, by improving the problemsolving-abilities of the students.

The course is organised in two modules. Every week there is a lecture (50 minutes) concerning the theory of Mechanics. This theory is applied to problems which serve as an example of the problems the students should be able to solve. Problemsolving can be trained by the students by making problems at home, and by attending working-classes (100 minutes a week). Every module lasts 10 weeks.

Rationale

In the working-classes the solution of the problems were shown by the teacher. More or less in a question-answer setting. The preliminary examinations showed many insufficient results. This was presumed to be caused by an insufficient attitude towards problem solving. When students got stuck in the problemsolving-proces, they often did not know how to overcome this situation. Two aspects seemed to be important. First of all, the students lacked confidence in their problemsolving abilities. Secondly, the students lacked a systematic approach. So their training situation should be optimised.

Theoretical framework

The theory about the discussion of example-problems in working-classes (Ten Dam & Smuling, 1987) and SPA (Systematic Problem Approach) (Smuling, Brants & Pilot, 1990) has been used as a cognitive framework. This lead to the selection of problems ranging from very easy to preliminary examination-level. In working classes, feedback was given instantly when students got stuck. By questioning, students were invited to elaborate on their problem solving process and the obstacles they meet. Feedback was further more given by the distribution of the elaboration of all of the problems.

The ideas of Freudenthal (1991) concerning meaningful learning lead to the selection of 'realistic' problems, problems chosen from the environment of the student (a bookshelf with books, etc.). Freudenthal introduced the idea of 'reinvention'. He meant that a student has to construct his knowledge-base by

1161

himself. He does so by 'mathematising' and solving realistic problems. Realistic problems are problems that do have meaning to the student. They do not have to be real, symbolic problems can be realistic if they are familiar to the student. If the course contains a correct order of realistic problems, than is the student able to (re)construct his knowledge-base.

Question
Does this selection and sequencing of problems result in the situation that more students then did in previous years master the ability to solve problems in Mechanics?

Method
Design
The research was carried out in a natural setting. A simultaneous experimental- vs controlgroup design could not be realised. Studentgroups from previous years were taken as controlgroups. The preliminary examinations of the different groups were different. But these tests correspond with the subject matter and the exercises treated in the lectures. The grade of difficulty should be the same.

Instruments
The instruments were the results on the preliminary examinations and a questionnaire.

Analyses
The data from the preliminary examinations will be analysed with a t-test. The data from the questionnaire will be analysed quantitively.

Result
The results of the group of '91-'92 were better ($\underline{M} = 56.3$) then the results of the group of '90-'91 ($\underline{M} = 53.6$), $t(131) = .93$, $p < .18$. This difference tends to significance.
This result is supported by the outcome of questioning students. Most students judged the sequencing of the exercises from easy to difficult positive. Most of the students who could compare this year courses with those from the previous years judged the improvement positive. The more realistic approach was mostly judged positive.

Discussion
The set-up of the experiment was quasi-experimental. Care has to be taken concerning the interpretation of the results. The results of the preliminary examinations and the questionnaire point both in the same direction: the selected problems and the sequence in which they were put seems to have the desired result. They will be put in an definitive form. As a next step, the materials for the lectures will be renewed. The possibilities of computer-presentations will be examined. The following step will be the implementation in an interactive computerprogram. Further more, the possibilities of (numerical) simulation of dynamic systems will be examined.

References

Dam, G.T.M. ten, & Smuling, E.B. (1987). Handleiding voor het begeleiden van werkcolleges. OC-Rapport nr. 62. Universiteit Twente, Enschede.

Freudenthal, H. (1991). Revisiting mathematics education. Kluwer Academic Publishers, Dordrecht.

Hout-Wolters van, B. Jongepier, P., &Pilot, A. (1983). Studiemethoden. Hoger Onderwijsreeks, Groningen: Wolters-Noordhoff BV, ISBN 90 01 41190 8.

Smuling, E.B., Brants, J., & Pilot, A. (1990). Oriëntatie op leren en onderwijs. Hoger Onderwijsreeks, Groningen: Wolters-Noordhoff BV, ISBN 90 01 80110 2.

1163

LEARNING AID FOR PROBLEM SOLVING IN PHYSICS

Hans van der Stam, Educational Development Centre, Hogeschool Enschede, Enschede, The Netherlands

Introduction

As there is a difference in the level of knowledge of a HAVO pupil (HAVO = general secondary education, higher level) in comparison with an HTO student (HTO = education in engineering on a practical base, polytechnics/colleges), the transfer from HAVO to HTO causes problems for many students.
This problem is recognized by the educational development centre of the Faculty of Engineering of the Hogeschool Enschede. In 1989 this centre started a project known as "aansluitingsproject HAVO/HTO" supervised by R. van Asselt.
In this paper a group of prospective students will be followed in their efforts to learn how to handle problems in physics in a systematic way. In the months preceding their exams they are instructed how to use the so-called SPA-method for a period of twelve weeks. There will be a midterm and a final examination; the results will be published in June 1992.

Objectives

A pupil from the HAVO is capable of solving problems in physics and chemistry when the problems are similar to the ones illustrated by the teacher. But when the problem is hidden in a large number of words and data and the phrasing of the question is not quite clear the pupil does not know where to start and just tries to find the answer by trial and error.
Although the SPA-method requires the student to use a specific structure which has to be followed through consistently, it is not meant to be a trick. On the contrary, the right knowledge of physics must be present and understood. Help is offered only with sorting and analysing data in a systematic way.

Theoretical framework

SPA stands for "Sytematische Probleemaanpak" (systematic problem solving). It originates from a doctoral thesis written by Mettes and Pilot (1980). Starting-point of Mettes and Pilot are learning theories developed by Galperin (1972) in combination with Talyzina's psychological bases of programmed instruction (1973) and the critical suggestions of Landa (1975).
The basic viewpoint of this thesis is that learning how to solve problems can be characterised by transforming the problem into recognisable standard problems.

1164

The SPA-method is based on the recognition of four stages when solving problems:

Stage I: Problem Analysis
Understanding the problem
- sorting data
- writing down the question
- estimating the answer

Stage II: Rearranging the problem
Designing a scheme for solving the problem
- general formulas
- specific formulas
- setting up the scheme

Stage III: Enforcement
Carrying out the scheme
- setting up the final relation
- carrying out calculations

Stage IV: Evaluation
Checking the results
- comparing the answer to the estimated answer
- checking the right symbols/units
- checking the significance of the solution

Experimental set-up
In the final year at the HAVO, the SPA-course is offered. When the pupils have understood the systematic approach of the SPA-method, they are asked to solve physics problems in accordance with the technique described above. For this they work individually or in small groups, using a "*Practice Book of Physics*" developed for this purpose. The physics problems were chosen from former HAVO exam papers and practice books of physics, and were rewritten in such a way that the SPA-method can be used for them.

Methods
During twelve weeks of two periods each the pupils are guided in the use of the *Practice Book* by a teacher (of their own school or of Hogeschool Enschede). The seminar in its purest form is used as a teaching method. This means that there is hardly any formal lecturing (only in case of problems for a group as a whole). The pupils are guided individually in using the four stages for a specific problem. When the pupil thinks s/he has found the solution, s/he can consult and take home a model solution to the problem concerned. In the year 1991/1992 210 pupils at 23 sites in the Twente and Achterhoek districts are offered the SPA-course and guided in this way.

1165

Results

Twice during the period of 12 weeks (in week 7 and in week 12) the group is tested to check understanding and application of the SPA-method. Like in the seminars, physics problems are used in this test, which is clearly aimed at evaluating the SPA-method. In other words, in checking the results marks are given for the correct use of the four stages. Therefore having just the (correct) answer will never score 100 %. The two tests are compared to each other to see if there has been an improvement in problem solving. As mentioned before, the main purpose of this project is to teach an improved strategy when solving physics problems. If the student shows that s/he approaches the problem in a systematic way, this will be regarded as an advantage. A correct use of stage I and II will therefore be valued heavily, whereas a correct use of calculation will be seen as standard and therefore valued as less important. With a view to the results it can already be concluded that pupils are not used to this way of dealing with problems. They easily relapse into their own trial and error methods.The results are not known at this moment, because the project has not been finished yet, but will be published during the European Conference on Educational Research.

References

Mettes, C.T.C.W., & Gerritsma, J.G.M., (1986). Probleem oplossen. Utrecht : Spectrum.

Mettes, C.T.C.W., & Pilot, A., (1980). Over het leren oplossen van natuurwetenschappelijke problemen. Enschede: TH Twente.

Smuling, E.B., Brandts, J., & Pilot, A., (1990). Oriëntatie op leren en onderwijs. Groningen: Wolters-Noordhoff.

THE IMPLEMENTATION OF THE "FACHHOCHSCHULE" IN THE NEW FEDERAL STATES OF GERMANY

Heidrun Jahn, Projektgruppe Hochschulforschung Berlin, Berlin, Germany

One result of the unification of Germany is the transfer and implementation of the West-German system of higher education in the area of the former GDR, which now comprises 5 new federal states (Bundesländer): Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt und Thüringen and applies also to East-Berlin, which is connected with West-Berlin now, building together the 16th independent state. Recommendations of the German Science Council (Wissenschaftsrat) together with the Unification Treaty have forced the new states to establish the "binary system" of Universities and "Fachhochschulen" as the unified frame of HE in Germany. In the GDR the "Fachhochschule"-type of higher education was not developed. The GDR had favoured - besides the traditional types of universities - special Institutes of Technology (Ingenieurhochschulen) and Technical Universities and - on a lower level - practically orientated and highly specialized technical and regional colleges (Fachschulen). The main reason for the recommended and already started foundation of 21 new "Fachhochschulen" for more than 50.000 students is the obvious success of this type of tertiary level education in the former FRG during its past two decades of existence. In detail the following aspects are usually mentioned:

- application orientated and scientifically based education;
- short and effective courses of study of three or four years;
- high rates of successful graduates;
- broad access to the various courses;
- cost-efficient type of higher education;
- regional importance and impact.

As against this some problems have to be considered which already characterized the situation of the "Fachhochschulen" in the former FRG, for example the increasing "numerus clausus", an unsatisfying structure and growth of the teaching and research staff and also the high teaching load of 18 hours per week for every individual professor. Other complaints and demands comprise:

- the access to research and its funding;
- the necessity for additional profiles and courses of study;
- the links between Universities and "Fachhochschulen";
- the equalization of salaries and income of both types of graduates namely in the public services.

1167

The main problem of the implementation of "Fachhochschulen" in the new states has to be seen in the fact, that no theoretical concept based the intended changes even though the already existing problems of these HE-institutions in the former FRG, especially the crucial relations between Universities and "Fachhochschulen", indicate that an elaborated and theoretical concept is urgently needed. The implementation is determined by the the political decisions within the Unification Treaty which principally lay down the takeover of the HE-system of the FRG into the new states. One main reason was the idea of the necessary equalization of law. Nevertheless expectations have been risen, in this case symbolized by the example of the "Fachhochschulen", that a new quality could be created and implemented based on the experiences of the old states, the merits of former institutions of HE in the GDR and the demands of an increasing and diverse expansion of higher education.

The Project-Team for Research into Higher Education (Projektgruppe Hochschulforschung Berlin) was founded in 1990 in order to document the on-going process of change and to assist and advise the governments and also the universities and other HE-institutions. As one issue of the research work the group has started to evaluate the implementation of "Fachhochschulen". The main questions to be answered are:

- To what extent are the recommendations of the Science Council accepted by the responsible governments and by the concerned staff and students?
- How is the realization of the recommendations set into practice under the concrete conditions in the new federal states?
- Are there already indications of promising further developments of the type "Fachhochschule" to be recognized?

The main characteristics of the research set up can be described as follows:
Central approach will be the comparison between the stated intentions and their realization at different subsequent stages (longitudinal study). The criteria will be derived from the Science Council recommendations.
Levels and panels of investigation will be the governments, staff, students and customers.
Methods of the research are:
Analyses of documents, questionnaires and interviews;
Information and feed-back to the involved and responsible groups and individuals about the results, problems and solutions by workshops and publications;
Special case-study of one "Fachhochhochschule" as for instance the new "Fachhochschule für Technik und Wirtschaft" in Berlin.
Recent results of the research from the initial phase reflecting the first half of 1992 will be reported at the ECER, mainly regarding the following issues:

- The implementation of the "Fachhochschule" in the new federal states shows different and contradictionary reactions: The staff of the former Institutions of Technology and Technical Universities, which shall be transformered into "Fachhochschulen" react with obvious resistance and complain of the downgrading of their institutions. Concerning these problems of acceptance, various reasons can be found: At first a lack of information regarding the merits of the "Fachhochschule", secondly the already reached status of the former GDR-Institutes including the access to research and the right to deliver the Ph.D.. A third reason for the resistance is the the irritating information about the problems of

international acceptance and recognition of the "Fachhochschule" and the lack of equality compared with universities. On the other hand one will find an active interest within the former engineering schools, where personnel expect the upgrading of their colleges and new perspectives for their work.

- At the beginning of 1992 a predominant agreement of the new governments to the recommendations of the Science Council can be considered. Difficulties regarding the realization emerge from the concrete conditions and constraints in the new states. The implementation turns out to be primarily a very complicated social process. One key problem is the question of staff for the new "Fachhochschulen". It is obvious that within the expected time scale neither a takeover of the existing research and teaching staff nor a major exchange or renewal of this staff will be possible. Thereby certain aims of the recommendations will presumably not be realized very quickly, especially the establishment of a diverse and attractive range of courses of high standard. This again may lower the interest and the enrollment of students in the new institutions.
- It can already be considered from the available data that despite of the predominant and comprehensive agreement the implementation of "Fachhochschulen" in the new states will lead to remarkable differences. The development of a new type of "Fachhochschule" will be restricted in the initial phase to only very few aspects. Some of these will be illustrated in detail at the conference.

EXPANDING HIGHER EDUCATION STRUCTURES: THE TYPE OF "FACHHOCHSCHULE" IN THE EUROPEAN CONTEXT

Günter Heitmann, Technische Universität Berlin, Berlin, Germany

Actual problems with the development of higher education in Germany and the explicit necessity to reflect the consequences of the European single market have resulted in an intensified debate of international issues and solutions, namely the mutual recognition of HE diplomas, international mobility and cooperation, quality assessment, duration of studies and last but not least the structural and funding aspects of an expanding HE system. An increasing demand for comparative research can be stated in order to analyse and to solve arising problems.
The "Fachhochschulen" in this European context have suffered up to now from a contradictionary role. On one side they still need to fight for their full recognition and on the other side they are well reputated and successful, not only from a national and industrialists point of view but also as a model for the internationally expanding structures of the tertiary level. What are the reasons for the mentioned difficulties and how can the future perspective be described?
On the background of various research activities three topics will be illustrated regarding past, recent and future conditions of the development of "Fachhochschulen":
The international problems of "Fachhochschulen" are to a great extent political and historical ones, only to a minor extent are they caused by the actual educational practice. The foundation of the "Fachhochschule" in 1970 was already a response to obvious deficiencies of the predecessors, the engineering schools, in the range of the European Communities policy. The main problem was the envisaged discrimination of 70% of the German engineers, who, from a European point of view were only assessed as technicians. Insufficiencies were stated concerning the status of the schools, the required entrance qualifications, the short duration of studies, the graduation. With the creation of the "Fachhochschule" most of these deficiencies were abandoned, which will be shown in detail in my contribution. In the European decision making process these developments were internationally acknowledged by the general directive of 1988 regarding the mutual recognition of HE diplomas. Early comparative research into the structures of the European educational systems have had a remarkable influence on the political process and decisions.

Continuing difficulties with an adequate evaluation and recognition of the "Fachhochschule" mainly derive from systematic differences concerning professional education and the diverse organization of the relationship between the labour market for highly qualified manpower and the HE systems. This hypothesis will be illustrated by the comparison of the HE systems of France, Great Britain and Germany. Despite the impression that these three countries are characterized by some kind of a "binary system" of HE (Universities - Grandes Ecoles, Universities - Polytechnics, Universities - Fachhochschulen) and that therefore convergency and harmonization is emerging, it will be pointed out that the three systems represent completely different responses to the needs of professional qualification and labour market demands.

This situation constitutes certain problems for the determination of equivalence and the promotion of international mobility especially for the "Fachhochschulen", and this can be shown by experiences with the ECTS-program of ERASMUS. The "Fachhochschulen" have successfully tried to find international partnership in the educational field by creating, for example, double-degree programmes on a bi- or multilateral basis. The desired expansion of research cooperation is limited by the mainly educational mission of the "Fachhochschule" and the poor funding of the "Fachhochschule", which is seriously criticised by the academic staff. This leads to the question of future developments and the third topic.

The prosperity of the "Fachhochschulen" in Germany and in the European context depends to a great deal on the development of HE structures in Europe. Obviously the demand for higher and continuing education will increase further. The shortage of public funding will force the search for cost-effective solutions of mass HE, which leads to the expansion of short-cycle, often practically orientated courses of study, as recent developments in Italy and France and also earlier ones in Spain and Greece, indicate. The further role of the "Fachhochschule" in Europe is, besides other aspects, influenced by possible future structures of HE. The following options and models will be discussed:

- a highly competitive model with a great diversity of independent universities, colleges and institutions;
- a structured "binary system";
- a mixed competitive model with a top tier (ivy league) and a diverse, possibly regionally orientated second division of mass higher education.

Research into HE can support the process of internationalization, mutual understanding and the development of the tertiary level for the use of students and the welfare of society. This contribution is based on an explorative and comparative study, especially in the field of engineering education in Europe, which shall lead to new concepts for the internationalization of engineering curricula and the increase of international cooperation and student mobility. Some possible examples of modular, common or integrated curricula and of mutual recognition will be pointed out. The research is not focused on quality evaluation and ranking procedures even though results of comparative research can be applied to these topics as well.

Through various EC-programmes, especially in the frame of Comett, research on the question of expanding structures of higher education and the development of new HE-institutions and courses of study influences the process of change and reorganisation of tertiary level education in the East-European countries. It seems that the type of "Fachhochschule" could be a useful contribution to the refereed process of reorganisation but does by no means represent the predominant strategy or institution of expanding the structures of higher education.

1171

BEST COPY AVAILABLE

UNIVERSITY STUDENTS' PERCEPTIONS OF THEIR EXPERIENCE OF TUTORIAL GROUPS

Charles Anderson, Department of Education, University of Edinburgh, Scotland, United Kingdom

Introduction

Tutorials, small group discussions which focus on some aspect of an academic discipline and which are customarily lead by a lecturer, still form an integral part of the teaching which undergraduate students receive in UK institutions of higher education. During 1991 and 1992 the author has been engaged in a study of the nature and quality of discussion in tutorials within the Faculty of Social Sciences at the University of Edinburgh. This study has looked at a number of tutorials conducted by each of ten tutors, from a range of very different disciplines within the faculty, and covering all the four years of undergraduate teaching. The principal methods of investigation have been:

- observation of the tutorials themselves,
- analysis of audio-recordings of the discussion,
- interviews with the tutors concerning their aims for, beliefs and feelings about tutorials,
- interviews with a sample of the students who took part in the tutorials which were observed.

This presentation focuses on describing the nature of the interviews which were conducted with individual students on their experience of tutorials, presenting findings from these interviews and considering what implications these findings may have for theoretical accounts of the negotiation of understanding in educational settings.

Method

The characteristics of the sample of the fifty students, drawn from all four years of undergraduate teaching, who were interviewed are described. The areas covered in the semi-structured interviews are then presented, followed by a detailed description of the style and techniques of interviewing that were employed. A rationale is also provided for adopting a style of interviewing which attempted to make transparent the interviewer's own purposes in asking particular questions, and which involved considerable use of requests for clarification and for expansion.

All of the students who were interviewed were also asked to complete an *Approaches to Studying Inventory* (Entwistle and Tait,1991) which provided data concerning their general purposes in studying and their style and approach to learning.
The methods employed to identify themes and to establish categories of description are described, along with the procedures adopted to ensure that a rigourous qualitative analysis of the interview data was achieved.

Results

The findings from the interviews are presented in two main sections. In the first section common reactions to specific aspects of tutorials are described and important differences in perceptions of these specific aspects are also highlighted. Topics which are discussed in this section include:

- general likes and dislikes concerning tutorials,
- features of tutorials which are described as helping or hindering participation in discussion,
- expectations and preferences for the role that a tutor should play,
- comments on tutors' style of questioning,
- likes and dislikes concerning other students' contributions to tutorials,
- perceptions of connections between tutorials and the rest of the teaching of a course,
- opportunities which tutorials provide to clarify ideas, and to explore problems in understanding,
- features which make active listening in tutorials easy or difficult,
- preferences for a tight, linear structuring of the discussion or for a more free-ranging discussion,
- perceptions of the utility of tutorial/seminar papers given by students, and descriptions of the preparation of student papers,
- differences in willingness to challenge interpretations made by other students and to engage them in debate.

The second section of the presentation moves from a consideration of specific topics to look at differences between students in their perceptions of the general purposes of tutorials and in what they see as the principal benefits of tutorials.

Discussion

The general concluding discussion centres first on a methodological point. Mention is made of how some of the findings from the student interviews illuminate, and facilitate the valid interpretation of, the results of the analysis of the transcripts of the tutorials themselves. It is argued that an analysis of talk and action within a particular setting needs to take into account the effects of the perceptions and interpretations which guide the contributions of the individual participants.
After looking at this and a number of other related methodological points, there is a short discussion of recent work guided by a 'neo-Vygotskian' perspective which has emphasized how gaining particular new meanings or a wider understanding of a topic involves a process of active negotiation (Anderson, 1991; Bruner, 1986; Wertsch, 1991). The findings of the current study are then used to illustrate how, at least within the context of group discussions in higher education, the process of

negotiating a new understanding of a topic may be powerfully mediated by the intentions and perceptions of the individual participants. Student perceptions and tutor actions which may act as barriers to the negotiation of new meanings are identified, and ways of dealing with these barriers are explored.

References

Anderson, C. (1991). Negotiating Understanding. Paper presented at EARLI conference, Turku, Finland, August, 1991.

Bruner, J. (1986). Actual Minds, Possible Worlds. Harvard University Press, Cambridge, Mass.

Entwistle, N. and Tait, H. (1991). The Approaches to Studying Inventory. Centre for Research on Learning and Instruction, Department of Education, University of Edinburgh.

Wertsch, J.V. (1991). Voices of the Mind: A Sociocultural Approach to Mediated Action. Harvard University Press, Cambridge, Mass.

STUDENTS' LEARNING ACTIVITIES BEYOND TEACHERS' OBJECTIVES

Diana Dolmans, Wim Gijselaers, Henk Schmidt, Department of Educational Research; Syb van der Meer, Department of Pediatrics, University of Limburg, Maastricht, The Netherlands

Introduction

Modern theories in the science of learning and cognition emphasize the importance of active learning. Many investigations indicated that competence is fostered not primarily by teaching to deliver knowledge, but through teaching to engender specific kinds of cognitive activity (Glaser, 1991). These results have led to the development of programs that aim to foster students' ability to become architects of their own knowledge and encourage them to eventually take full responsibility for their own learning. Such approaches are quite different from teacher–centred approaches in which teachers determine what information should be learned, how it is to be learned, and in what sequence by dispensing information in lectures. In problem–based learning by contrast, students actively engage in acquiring knowledge and define themselves to a large extent the content to be mastered. Problem–based learning is assumed to be an instructional approach reflecting scientific conceptions of the conditions of human cognition and learning. Previous to arguing this statement, the process of problem–based learning will be shortly described.

The process of problem–based learning starts with a problem consisting of a description of a set of phenomena in need of some kind of explanation. A group of 8 to 10 students discusses these problems and tries to explain the phenomena in terms of underlying processes, principles or mechanisms (Schmidt, 1983). During this discussion issues emerge needing further exploration. These student–generated learning issues are topics that each tutorial group decides are prerequisites to a better understanding of the problem under discussion (Blumberg, Michael & Zeitz, 1990). As such, learning issues serve as the starting–point for students' learning activities and as a guide for studying literature relevant to the problem at hand.

Barrows and Tamblyn (1980) claim that problem–based learning has several advantages. First, students do 'learn to learn', so that they can make their learning relevant to their own educational needs. Second, the student–centred learning approach increases motivation because students themselves define their own

learning issues and decide what is relevant for their learning. Third, students learn to see gaps in their own knowledge and learn to evaluate their own strengths and weaknesses. Finally, because students conduct the literature search themselves, students learn to find the material on their own and acquire the ability to continue their education after they graduate. In other words, students learn to reflect upon and control their own learning activities and develop self–regulatory skills (Glaser, 1991). Problem–based learning is an instructional approach in which students must be active rather than passive learners complying with modern theories in research on the science of learning and cognition.

Student learning in a problem–based curriculum is assumed to be dependent on the nature of the problems presented. While constructing problems teachers have certain objectives in mind. These faculty objectives define the subject–matter that students need to master and suggest which learning issues should be generated by students. However, since problem–based learning accords students an active role in learning, students themselves more or less define the content to be mastered. This self–directedness allows students to generate learning issues beyond faculty objectives. These are learning issues not specifically required to be generated and not corresponding with the preset faculty objectives. Nevertheless, these learning issues, although unexpected, may still fit within the framework of the problems at hand. The study to be presented deals with these unexpected learning issues. Leading questions of this study are:

- To what degree do students generate unexpected learning issues?
- Can these learning issues be considered relevant to the course as a whole or do they lead to irrelevant learning activities?
- Why do students generate these learning issues?

Method

Subjects

This study was conducted at the Medical School of the University of Limburg, The Netherlands. The first four years of the problem–based curriculum are organized in units, a series of 6–week courses.

Materials

The data were collected during the sixth unit of the second year, in the 1990–91 academic year, containing 12 problems concerning topics such as childbirth, psycho–motor development and psychosexual development. Teachers had 51 objectives in mind while developing the 12 problems. The average number of faculty objectives for each problem was 4.3. Both students and tutors were randomly assigned to 12 tutorial groups. During this unit learning issues generated by the students were collected. Each tutor was asked to record the learning issues generated by his or her group for each problem. Tutorial groups generated an average of 3.6 learning issues for each problem. These learning issues were collected by the researchers at the end of the unit.

Procedure

A few weeks after the end of the unit, raters were asked to judge the correspondence between faculty objectives and student–generated learning issues. The raters were students who participated in the tutorial groups of this unit.

Pilot–studies suggested that results from teacher–raters and student–raters were quite similar. The raters had to judge whether a particular learning issue: (1) definitely corresponds with a particular faculty objective or (2) definitely does not correspond with that objective.

Analysis

If both raters agreed that one or more learning issues corresponded to one faculty objective, then this particular faculty objective was judged as 'identified by the students'. If both raters agreed that the faculty objective is definitely not identified, then it is assumed by the researchers that the students failed to study the content domain specified by this objective. If both raters agreed that a learning issue did not correspond to any of the faculty objectives, then this student–generated learning issues was assumed to be unexpected. The focus of this study is on these unexpected learning issues. In order to analyze whether these learning issues guide students to content–domains which are still relevant in relation to the course content, a teacher who was familiar with the unit content, was asked to judge the relevance of these learning issues. He had to indicate whether each issue was:

- irrelevant;
- fairly irrelevant;
- neutral;
- fairly relevant ;
- relevant in relation to the unit content.

Furthermore, in order to explore why students generated these unexpected learning issues, an unstructured sorting procedure was conducted in which each learning issue was printed on an index–card and the teacher was asked to sort these learning issues into meaningful categories.

The interrater–agreement for each pair of raters, varied between 67.5 percent and 97.3 percent, with an average of 78.9 percent. Since these percentages are overestimated, because of agreement based on chance, kappa–coefficients were computed for each pair of raters. The kappa–coefficient takes into account the expected proportion of agreement that would occur if assignments were made by chance alone. The kappa–coefficient for the pairs of observers, varied between .23 and .64, with an average of .45. The average percentage of .45 demonstrates a moderate agreement (Landis & Koch, 1977). These results seem to support the reliability of the match procedure.

Results and discussion

The overlap between faculty objectives and student–generated learning issues for each problem varied between 27.7 and 76.3 percent and differs across problems ($F(11,143)=6.84$, $p<.000$). The average percentage of overlap for the 12 problems together was 64.2 percent. The 12 tutorial groups generated 520 learning issues for the 12 problems. The number of student–generated learning issues that did not correspond with faculty objectives was 32 (6.2 percent). Teacher's judgments about the relevance of these learning issues revealed that 47 percent of these learning issues were rated as 'relevant' or 'fairly relevant', 28 percent were rated as 'neutral' and 25 percent of these faculty objectives were rated as 'fairly irrelevant' or 'irrelevant'.

The sorting procedure which was conducted to identify key features of these unexpected learning issues revealed four categories. Learning issues in the first category were associated with extra curricular activities such as skills training. These issues are probably generated because students did not attend the extra curricular activities or because they did not really understand a specific issue addressed during a training. The second category involves learning issues which are not yet intended by the teachers who developed this unit, since these issues are dealt with in later years of the study. The first two years of the Maastricht medical school emphasize normal physiological and pathophysiological processes, patient management and medical intervention are addressed to the students in the third and fourth year. The third category contains learning issues associated with gaps in students' prior knowledge. Although these learning issues are related to the course content, teachers expected that students would pay minor attention to these issues because they are only a prerequisite to understand the problem at hand and these issues were assumed to be addressed in earlier years. The fourth category contains learning issues which can only partially be derived from the content of the unit because they are connected with details in the text. These learning issues arose out of students' personal interest or experiences and are mostly associated with topics often stated on television, newspapers and other mass media.

Although problems are the starting point for students' learning activities and determine the generation of learning issues, it should be noted that learning issues are only a possible reflection of learning, implying that it is not sure that students actually pursue these learning issues in all cases. Students will not employ certain learning activities related to a particular problem when they have already studied this material before. Students may decide it to be more relevant to spend time on other issues. On the other hand, it seems clear that during independent self–study and discussion students frequently learn about other related topics that go beyond student–generated learning issues.

In summary, analysis carried out in this study suggest that problem–based learning seems to foster students' ability to become architects of their own knowledge. Students seem to be able to eventually take full responsibility for their own learning, since they generate learning issues that, although unexpected, were judged as relevant and worthwhile. Moreover, these issues mostly seem to be arisen from students' personal interest in the subject–matter or lack of prior knowledge indicating that students use the freedom to choose learning issues to modify their knowledge deficiencies.

References

Barrows, H. S., & Tamblyn, R. M., (1980). Problem–based learning. An approach to medical education. New York: Springer.

Blumberg, Ph., Michael, J. A., & Zeitz, H. (1990). Roles of student–generated learning issues in problem–based learning. Teaching and Learning in Medicine, 2, 149–154.

Glaser, R. (1991). The maturing of the relationship between the science of learning and cognition and educational practice. Learning and Instruction, 1, 129–144.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33, 159–74.
Schmidt, H. G. (1983). Problem–based learning: Rationale and description. Medical Education, 17, 11–16.

1179

DIRECTING LEARNING ACTIVITIES BY MEANS OF TASKS IN PROBLEM BASED LEARNING

K. van Wijngaarden, IOWO, & J. Willems, Katholieke Universiteit Nijmegen, Nijmegen, The Netherlands

The principle idea behind problem based learning (PBL) is that learning should be organized around problems which are related to the profession, rather than around subjects which are derived from academic disciplines. Problems and other kinds of tasks are the starting-point for students learning activities. The students themselves determine their learning goals and the learning activities they will undertake (Schmidt, 1982). Because education is a goaldirected activity, the curriculum tries to direct the students learning activities by giving them different kinds of assignments. Still unknown is whether the different kinds of tasks lead to the intended learning activities. It is also not clear which elements of a task contribute to the directing effects of those tasks (if any). To get a better understanding of the effects of tasks in PBL in directing the students learning activities, two related research studies were performed.

The first study was exploring in nature and centred around two questions. First: do students undertake different learning activities when performing different tasks? Second: to what extent do different elements of tasks have a guidance effect on actual learning activities?

The research took place in the second block of the curriculum that was offered in the first year of the faculty of Policy Studies in Nijmegen. This curriculum is organized in accordance with the PBL-philosophy. Next to a number of rather traditionally structured program-elements, there is a subdivision into five PBL blocks concerning specific themes on several policy-areas. The tasks that students could undertake were formulated around these themes. The research took place in the second block of the curriculum that was offered in the first year of the faculty of Policy Studies in Nijmegen.

In this research project 120 students (10 groups of 12 students each) volunteered as subjects. Each time the students completed a task, they were given a questionnaire about their learning activities while performing that task (actual learning activities: ALA). The questionnaire used was based on the ILP (Schmeck e.a., 1983) and adapted for use in a PBL setting. Nine tasks were involved in this research: 3 problem tasks, 3 study tasks and 3 application tasks.

A principle component analysis of the questionnaire resulted in two stable factors: deep processing (α = .64) and structuring (α = .62). These factors explained 55% of the variance.

As a central point in this research it was examined whether the learning activities differed with each kind of task. This was tested by a t-test for paired examinations. As a result it appeared that there was more structuring at problem tasks (t = 4.65) and study tasks (t = 4.50) than at application tasks. At application tasks (t = 3.47) and study tasks (t = 3.73) there was more deep processing than at problem tasks. So contrary to the expectations most learning activities (structuring and deep processing) were found at study tasks. Problem tasks did not result in more deep processing activities.

It was also examined whether certain taskelements could be distinguished in these tasks which contribute to the intended effect of the different kinds of tasks. Seven elements were distinguished, varying from an introduction at the tasks to a specific instruction to carry-out the task.

For each element it was tested whether the presence lead to more deep processing and more structuring activities by means of a t-test for paired observations. It appeared that the presence of the taskgoals, relevant literature and specific instructions resulted in less structuring activities of the students and - to a certain extent - more deep processing activities.

It was concluded that especially in problem tasks more guidance was needed to get the intended learning activities. In particular were these kinds of tasks in need of a well-described task goal and an instruction to carry-out these tasks.

In research project 11 two groups of 65 students each were involved. The setting for this research was again a block of the regular program. This research concerned 3 problem tasks of that block. One group received the tasks as developed by the faculty (standard tasks). The tasks for the other group were adapted based on the results of research I (adapted tasks), i.e. a task goal and a specific instruction were added.

The main questions concerned:

- the difference between the two groups in planned and actual learning activities;
- the resemblance between intended and actual learning activities;
- the influence of the learning style on the intended and actual learning activities and the extent to which this influence differed in both groups.

Preceding each task every student had to answer a questionnaire concerning the way he intended to perform the task (planned learning activities). After performing the task the students had to answer the questionnaire concerning their actual learning activities while performing on the task. In addition the students learning styles were measured with a condensed version of the Inventory of learning styles (Vermunt, 1991). It was assumed that the learning activities and the way students react at the regulation attempts by the instruction depend on the learning style of the students. To answer the questions in this research the data were analyzed with a t-test to test the differences between the two groups and a MANOVA to examine the influence of the other variables.

The results can be summarized as follows. There is no difference between the two groups on the intended learning activities nor on the actual learning activities. There are indications that there is a relation between the learning set, intended learning activities and actual learning activities. In addition it appeared that the learning style influenced the effect of the tasks on the learning process.

References

Schmeck, R.R. (1983). Learning styles of college students. In R. Dillon & P.R. Schmeck (eds.), Individual differences in cognition, 1. New York: Academic Press

Schmidt, H.G. (1982). Activatie van voorkennis, intrinsieke motivatie en de verwerking van tekst, Apeldoorn: Van Walraven

Vermunt, J.D.H.M. (1991). Leerstrategieën van studenten in een zelf-instructieleeromgeving, Pedagogische Studieën, 68, 7, pp. 315-325

PREPARING STUDENTS FOR CLINICAL PROBLEM SOLVING IN THE PRE–CLINICAL CURRICULUM

Theodorus Jan ten Cate, & Egbert Schadé, Faculty of Medicine, University of Amsterdam, Amsterdam, The Netherlands

Introduction

The primary objective of the training of medical students may be the competence to solve medical problems. Yet, many students have difficulty handling patient problems when they enter the clinical ward, even after years of medical education. The issue is: how can students be better prepared for clinical problem solving? Designing adequate instructional procedures that lead to this objective requires both knowledge of clinical problem solving and knowledge of effective instructional methods.

Description of the goal: essentials of clinical problem solving

The competence of solving clinical problems, to be defined as a cyclical reasoning process of generating hypotheses, inquiry and testing (2), appears to be rather content–specific. Being able to solve problem A does not imply the ability to solve B in the same content–domain (18,23,34). Also, the ability to solve clinical problems may to a large extent depend on experience in patient care. Experience with specific patients and their clinical context facilitates the retention of clinical information (10,13,31,32). A broad skill in clinical problem solving probably therefore requires broad a experience. Medical schools are indeed faced with a significant problem, since clearly practising as a doctor seems necessary to acquire the problem solving skill. Although little time is given, extensive clinical experience should somehow be realized in the curriculum.

Another consistent finding is that medical students starting the clerkship–period do not adequately use biomedical knowledge in problem solving (1,21). The existence of a three–stage model has been suggested: "In the first stage biomedical and clinical knowledge develop separately with an emphasis on biomedical knowledge. In the second stage, taking place in the clerckship period during which students are for the first time exposed to real–life clinical cases, emphasis is shifted towards the acquisition of clinical knowledge. The final stage is reached when both knowledge bases have matured to a sufficient extent, and compilation of biomedical knowledge and integration into clinical knowledge occurs"(6). A second task for the medical school arises: designing a curriculum in

1183

which efficient integration of biomedical knowledge within clinical knowledge is realized.

Search for instructional procedures leading to transfer

The quest for instructional procedures that lead to broad, stable problem solving skills, with as wide transfer as possible, is of course not characteristic for medical education alone. Lack of transfer outside the content–domain is often found with the teaching of cognitive skills; transfer within the content–domain may or may not occur (9). How should the intermediate–distance transfer called 'problem solving' (14) be best trained? A distinction has been made between a 'low road' (automatic) transfer and a 'high road' transfer, i.e. resulting from a deliberate mindful abstraction from non–essential features of a presented example (24). Assuming that both types of transfer are useful objectives of (medical) education, 'low road'–transfer requires practice and 'high road'–transfer requires mindful abstraction, for instance through verbalization of meaningful understanding, analogical reasoning etcetera. Being mindful or not may be viewed as a personality characteristic of students (24). Maybe one should try to induce a habit of mindfulness through giving attention to it in subsequent courses and to foster mindful transfer through methods that realize achievement motivation (14). Sufficient self–competence and achievement motivation is considered necessary for mindful transfer: 'unmotivated learners are likely to withdraw or proceed mindlessly' (14). Vermunt (33) advocates a process–oriented approach which may well turn out to stimulate the internal motivation, necessary for high road transfer that occurs in problem solving.

This short overview suggests using educational procedures in the medical curriculum that stress mindful abstraction through cognitive tasks (for high road transfer), frequent practice (for low road transfer) and an educational setting that stimulates self–regulation of learning.

Current approaches in medical education

Many recent transitions medical curricula (12,29) in structure, content, instructional methods and testing methods are somehow linked to the wish to teach problem solving skills. Integration of basic science and clinical subjects into theme–wise entities, clinical demonstrations early in the curriculum, small–group education, training of clinical skills before entering clinical wards and the introduction of simulated patient techniques all hope to prepare students better for dealing with patient problems. Notably one approach has internationally served as a model: Problem Based Learning (PBL), designed by Barrows c.s. (3). Basically in this curriculum students are, from the very beginning, continuously confronted with problems, serving as the major stimuli to acquire knowledge. It is expected that students from these curricula will be better prepared for patient care, i.e. for solving (clinical) problems.

Since some years comparative studies of PBL versus conventional curricula appear in the literature (2,7,16,26,25,31,20,15). As far as factual knowledge is concerned, alumni from non–PBL schools seem to score at the same or somewhat better level than alumni from PBL schools. When problem solving skills are compared, the results are inconclusive. Santos–Gomez et al. (25) find no difference in critical thinking and independent learning between PBL and non–

PBL educated residents. Patel c.s. finds that PBL–students show the type of systematic thinking that is taught: more backward reasoning, more systematic use of clinical information, more elaboration and generation of multiple diagnostic explanations through hypothetical–deductive reasoning. However, the PBL students also tend to make more errors in their explanations than non–PBL students (20).

There has undoubtedly made much progress in techniques of instructing problem solving skills by the concept of PBL. In some respect, however, PBL may possibly be less effective. Potential dangers of full integration of clinical and biomedical education have been pointed out (20,6) "A considerable amount of confusion may be avoided by ensuring that one of the two domains is highly overlearned before the other is encountered. In fact, the use of erroneous basic science explanations by the PBL–curriculum students () indicates that there is a genuine danger of such confusion in an integrated curriculum" (20). Indeed, PBL is primarily designed as a general didactical method and not a specific instructional tool to learn clinical problem solving; learning to solve problems is not the central objective in PBL (35). Expecting better solving of clinical problems should therefore not be justified only because of having gone through a PBL–curriculum, since problem solving skills are highly content–specific.

Requirements for a curriculum preparing for clinical problem solving

Resuming the above, the following suggestions can be made:

- build 'clinical' experience by confronting students with many, different and relevant problems;
- integrate biomedical and clinical teaching in such a way that the domains facilitate and not hinder each other;
- formulate cognitive tasks to foster mindful abstraction;
- realize frequent practice and rehearsal;
- stimulate self–regulation of learning.

In designing the medical curriculum at the University of Amsterdam (a cooperative effort with the Free University of Amsterdam) these requirements have tried to be met. The framework of the pre–clinical curriculum consists of a combination of (a) integrated blocks (80%) and (b) line–wise clinical teaching (20%) (apart from individual practicals (e.g.,nurse–aid period, research period).

(a) The blocks all have a theme–wise, objective–based, multidisciplinary content, in which basic and clinical sciences are combined. Central issues in these blocks are 'patho–biological' concepts with clinical relevance. Basic–science subjects are taught from the point of view of relevance for patho–biology. Clinical topics, i.e. diseases with their characteristics, are related to the patho–biological concepts and are systematically taught within the blocks. Blocks close with a formal examination of biomedical and (systematic) clinical knowledge.

(b) The line–wise clinical teaching is designed to train clinical problem solving. Every week one day (more in the higher years) is devoted to (i) clinical demonstrations and modeling of the problem solving by clinical teachers (ii) practising of cognitive skills in clinical problem solving in small groups and (iii) training of psychomotor and verbal skills. These three are increasingly integrated during the years. The point of departure in (i) and (ii) is not the systematic approach (cf. the blocks) but, invariably, some clinical manifestation of a patient–

1185

case. In dealing with this case, knowledge of prior block–themes is required. The line–wise clinical education in this way constantly applies prior systematically taught biomedical and clinical knowledge in clinical situations. Very little new content–information is presented, only when necessary to understand a specific case.

Specifically the small group training (ii) is designed to meet the educational requirements of active practice, building experience, exerting cognitive tasks and stimulation of self–regulation of learning. To this end a peer teaching format has been chosen. Groups of 12 students meet every three weeks during year 2 through 4. Each session 1–3 clinical cases are in discussion. Cases consist of a written introduction of the patient with his or her clinical manifestations and a series of questions (e.g., what would your hypothesis be? what patho–biological bases do you have for this opinion? what epidemiological reasons do you have to expect outcome X? etc.). Of each group 3 students operate as 'referents' and are provided subsidiary information on the case (such as results of laboratory tests). Referents are considered relative experts and to lead the session, including the discussion. They may be asked to give a mini–lecture during the case presentation on a prior assigned relevant topic, to help proceed the problem solving process. Each group has a permanent clinical teacher to guide the learning process and to assess the performance of the referents and – marginally – the other students. Active participation is required and each student is a referent once every 4 sessions. All students are expected to read assigned material and prepare answers to the questions. Furthermore, the cases are designed to incorporate knowledge of prior blocks making it necessary for students to regularly review material. Specifically the referent function may stimulate active elaboration of prior knowledge (28), which is considered beneficial for retention (27) and clinical thinking (8). This curriculum has started in 1990/1991. Our first evaluation data indicate that the method is well accepted by clinical teachers and students and that the training of clinical problem solving may be effectively realized within this format.

References are available on request from the first author

1186

EVALUATING TEACHING PERFORMANCE IN CANADIAN UNIVERSITIES

Alenoush Saroyan, & Janet G. Donald, Centre for University Teaching and Learning, McGill University, Canada

Objectives

The overall purpose of this study was to discern the policies and practices in Canadian universities which contribute to the quality of teaching and direct its assessment. The first specific objective was to examine the way in which Canadian universities define good teaching and the importance they give to the defining factors. The second objective was to determine the relative importance ascribed to teaching in comparison with other academic responsibilities. The third objective was to establish which methods are used most frequently to assess the quality of teaching and the procedures that are involved in their use. Large and small universities were used as the context within which trends and patterns were described.

Perspective

Three perspectives have guided this study. The first is the way in which good teaching has been characterized. Instruction has long been considered to play a pivotal role in students' academic development and has been credited with influencing their achievement, emotional well-being and value systems. Higher education will play an increasingly prominent role in determining a nation's competitiveness in the world marketplace and so will the quality of instruction. While accountability to students and to the community at large is growing in universities and colleges and there is greater attempt to ensure the delivery of the best education possible (Centra, 1980), the perspectives on teaching and the values assigned to them seem to vary from one institution to another. Some universities actively promote teaching and its improvement by means of supporting research into teaching and professional development activities for their faculty. Others are less willing to devote resources toward the improvement of teaching (Abrami, 1985). While there is increasing agreement that good teaching can be characterized in terms of attributes such as knowledge and enthusiasm for the subject matter, ability to communicate with students and to provide them with a conceptual framework of the discipline, ability to model thinking and the methods of inquiry associated with a field or discipline, and the ability to encourage students to actively participate in the teaching and learning process (Donald,1984; Seldin, 1980), there is still considerable variance among professors, university

1187

administrators and students in the level of importance ascribed to each characteristic. For example, university students often define good teaching in terms of personality traits, using words such as enthusiasm, rapport, charisma, dynamism, or warmth (Feldman, 1989). Others have described good teaching in terms of course presentation, the learning environment, and showing genuine interest in the student (Thomas, 1990). Thus, in order to assess good teaching, it is important to determine the degree of importance that universities ascribe to characteristics of good teaching.

The second perspective is the value of teaching in the evaluation of academic performance. In the university, there are three generally accepted areas in which quality indices can be discerned: research, teaching, and service. Although university policies tend to support excellence in all three areas, there are large variations by institution and by discipline in the importance given to each area and in the combination of duties which professors undertake. Teaching can be considered rejuvenating and at times inspiring but there is evidence that undergraduate teaching is not ranked highly by professors whereas teaching graduate students and conducting research are (Bess, 1982). The fact that the university provides knowledge and expertise to the surrounding community also affects the emphasis it puts on research or teaching. As both Clark (1987) and the OECD Report (1981) have indicated, pressure from outside groups, whether granting councils, government or industry, can affect university policy and practice regarding the importance given to teaching and research. When the question of the assessment of teaching is brought to the forefront, these realities cannot be ignored. This may be the time to seriously reexamine the mission of higher learning institutions and begin to value teaching not only in rhetoric but by implementing policies and practices which foster such an attitude.

The third perspective is the objective evaluation of teaching performance. In spite of twenty years of research on the evaluation of teaching in universities, reliable and valid methods for the evaluation of instruction are still being developed and tested. According to the literature, most North American universities support the evaluation process and have developed basic procedures for assessing faculty performance (Marsh, 1987). As a result of the interest in and concern about teaching, many universities have taken steps to establish policy, identify dimensions of effective instruction, and introduce methods to evaluate teaching. The degree with which the systematic evaluation of teaching has been incorporated into university policy and the congruence of such policy with actual practice is not very clear. Such information would seem to be fundamental to a comprehensive examination of teaching assessment.

Methods

The survey method was used for data collection. Participants received a questionnaire in advance and responded to it during a telephone interview. To identify the participants, initially, the office of the president of each university was contacted and the name of a designated representative was solicited. Most of the representatives named were vice presidents, and associate or assistant vice presidents. In all, fifty-four universities (88%) responded directly to the interview questionnaire. The primary source of data was the completed questionnaire which consisted of 22 questions in three formats: rating on a Likert scale, selecting

multiple responses, and selecting a single response. The SPSSx package was used for the analysis of data. Two procedures were used. First, a confirmatory factor analysis was carried out on items which were listed under some questions as definitions of a construct. The purpose of this analysis was to establish the correlation of each item with the emerging factor(s). Second, crosstabulations were run to generate frequencies and percentages of responses by university size. Size was determined by student enrollment: Small universities had less than 6600 students (the median) and large universities more than 6601. Twenty eight institutions were classified as large and 26 as small.

Results and Discussion

In each of the 4 questions for which the factor analysis procedure was used, over 80% of the variance was attributed to only one factor. The analysis yielded extremely high correlation values between each item and the identified factor; mostly .8 and .9. This confirmed the strength of the statements listed as indicators of good teaching, dimensions used in the assessment of good teaching, methods of assessment, and academic responsibilities.

Crosstabulations yielded the following results. With respect to the first objective which was to determine how Canadian universities define good teaching and what importance they give to different factors the most salient results were as follows: Four indicators, being up to date in subject matter, effective communication skills, defined grading procedures, and the ability to stimulate ideas and thinking were found to be highly important to the majority of the respondents. Large and small universities agreed on the importance of the same indicators. However, small universities also placed great importance on student centred teaching and the maintenance of office hours. Knowledge of graduation requirements was rated as unimportant by almost half of all the respondents. The second objective was to examine the relative importance ascribed to teaching in comparison with other academic responsibilities. Overall, the most important criterion was judged to be teaching competence, closely followed by research productivity. Other important criteria were academic degree and effective communication skills. Large and small universities had inverse ratings for teaching and research. Large institutions rated research productivity very important more frequently whereas small universities rated teaching competence as such. In addition, large universities also placed greater importance on academic degree and the supervision of graduate students while smaller institutions rated student centred teaching and service within the institution as highly important. The third objective was to investigate the methods by which teaching is assessed, the policies which guide their use, and the procedures for their administration. About half of the universities did not have a policy on the quality of teaching. This included one third of the large and over half of the small universities. It was found that most policies emphasize that the results of teaching assessment be used for the improvement of teaching and for tenure and promotion purposes. In practice, most universities use student rating forms to evaluate the quality of teaching. This is typically done by instruments designed by departments as opposed to one standard form used across the university. In this respect, small universities are more likely to use a single form than large universities. The procedures used for administering, compiling, and distributing the results of student rating questionnaires vary. Typically, they are administered in the

last two weeks of the term and by a student representative. Professors and department chairs receive the results in the form of computer printout or raw data. In the event that teaching is rated poorly, an interview with the chair ensues. The resources available to universities used to improve teaching include committees on teaching and learning at the university and department levels and a centre for teaching improvement.

Educational or scientific importance of the study

The findings of this study have multiple implications. This comprehensive dossier on attitudes toward teaching across Canadian universities can assist university faculty and administrations in redefining or clarifying university mission statements which deal with teaching. The study also contributes to the movement to elevate consciousness on the part of university administrations on the importance of recognizing and rewarding good teaching, by providing them with comprehensive and representative data.

References

Abrami, P. C. (1985). Dimensions of effective college instruction. The Review of Higher Education, 8(3), 221-228.

Bess, J. (1982). University organization: A matrix of the academic profession. New York: Human Sciences.

Centra, J. (1980). Determining faculty effectiveness. San Fransisco, CA: Jossey-Bass Inc.

Clark, B. R. (1987). Academic tribes and territories. Milton Keynes: Open University.

Donald, J. G.(1984). Quality indices for faculty evaluation. Assessment and Evaluation in Higher Education, 9(1), 41-52.

Feldman, K. A. (1989). Instructional effectiveness of college teachers judged by teachers themselves, current and former students, colleagues, administrator, and external (neutral) observers. Research in Higher Education, 30, 137-194.

Marsh, H. W. (1987). 'Students' evaluations of university teaching' research findings, methodological issues, and directions for further research. International Journal of Educational Research, 11(3).

OECD (1981). The future of university research. Paris: Organization for Economic Cooperation and Development.

Seldon, P. (1980). Successful faculty evaluation programs. Crugers, NY: Coventry.

Thomas, A. (1990, April). The basis of student judgement on teaching quality in different academic areas. Paper presented at the annual meeting of the American Educational Research Association, Boston, MA.

THE FIRST YEAR EXAMINATION AS NEGOTIATION; AN APPLICATION OF COLEMAN'S (1990) SOCIAL SYSTEM THEORY TO LAW EDUCATION DATA

Ben Wilbrink, Centre for Educational Research, University of Amsterdam, Amsterdam, The Netherlands

Law students at the University of Amsterdam in their first year of study typically invest only 800 hours in the first year examination, using the other 800 hours for leisure activities. Yet marking is not lenient: the attrition rate is very high, and a high proportion of 'unsatisfactory' marks is given in each of the six parts of the examination. Obviously there are some inefficiencies in this educational system. Obviously, because this situation repeats itself from year to year, individual actors involved do not have any influence on it, nor do policy makers. The question to be answered by the educational researcher is, given data on time expenditure and marks received, to describe this kind of educational system in terms that open up possibilities to manipulate the situation in the direction of higher efficiency. This kind of description is exactly what the paper will present, using a social system model as recently presented by Coleman (1990), and a data set obtained from the already mentioned first year examinations in law, in the years 1983-1989. (The methodological aspects and implications of the Coleman model are treated and discussed in a companion paper 'Modelling the connection between individual behaviour and macro-level outputs.').

The first year examination in Dutch higher education typically is subdivided in several tests, in the case of the study of law there being six tests, the first two scheduled after the end of the first trimester, the other four during the latter part of the third trimester. The examination result is determined by the results of the separate tests, the marks of every test must be satisfactory, some compensation of lower marks by higher marks being allowed. The Dutch scale for marks runs from 1 to 10, all marks below 6 being 'unsatisfactory', and a '10' being the best mark.

The actors in this social system are the teachers and the students. Students, by preparing for the tests, exchange time for marks. The teachers can be seen as exchanging marks for time spent. The exchange of time and marks is only indirect, the test itself functions as a clearing house for the exchange.

Now suppose there exists a general exchange rate of marks and time in this system; for example: to get an 11 % higher mark an arbitrary student has to invest or 'pay' 7 % of his time budget. Now students are free to invest as much or as little time as they wish, knowing perfectly well that spending more time on their study and less on leisure activities will result in higher expected marks. Time spent on leisure activities as well as marks received will give the student a certain

1191

satisfaction, the relative amount of satisfaction depends on the amount *controlled* (possessed or consumed) of the particular good (marks) or resource (time), and the *interest* the person has in the good or resource. These *interests* are a personal characteristic of the student. The same can be said of the other (type of) actor in this social system: the teacher. An important assumption on human behaviour, one that is routinely made in the field of micro-economics, is that an actor is less and less willing to exchange still more time for marks (in our case) the higher his or her expected mark gets. Now assuming that actors do *maximize* the satisfaction they derive from marks received and time spent in leisure, it is possible, given the *interests* of the students and the amount of time they have before any exchanges take place, to determine the time individual students will spend in preparation for their tests.

Interests can not be observed, but it is possible to work the other way around, from the available data on marks and time spent in preparation for the tests to the interests of the actors. Remember the assumption of a general exchange rate of marks and time. Under the existence of a general exchange rate, everybody pays the same price for a particular good, say marks, and gets the same price for, say, time in exchange of marks. The social system then functions like a market where none of the players individually is able to influence exchange rates. From observations on time and marks it is then possible to determine the exchange rate, or the relative *values* of marks and time. Table 1 gives the *values* of marks for each of the six tests, in each of six years. The *values* of time simply are 1 minus the value of marks.

test	1983/4	1984/5	1985/6	1986/7	1987/8	1988/9
1	0.36	0.39	0.40	0.38	0.40	0.38
2	0.42	0.39	0.39	0.35	0.34	0.37
3	0.43	0.43	0.46*	0.48*	0.51	0.48
4	0.41	0.43	0.41	0.44	0.38	0.41*
5	0.36	0.35	0.33	0.32	0.38	0.36
6	0.37	0.30	0.27	0.27	0.31	0.26
examination:	0.40	0.39	0.38	0.38	0.39	0.38

Table 1
The values of marks 1983/54 - 1988/9 for 6 tests, and the examination.

Notes: value of time spent = 1 - value of marks; N (students) = 109, 140, 88, 120, 106, 96 resp. over the years; only the students participating in all 6 tests; for three *-marked tests no questionnaires were received, the reported values were derived from estimations of the missing data.

The examination itself can be regarded as a social system in a more comprehensive analysis. Here the student has a budget of, say, 1440 hours, to be spent on each of the tests, or on other study activities, leisure activities or work. The values reported in the row marked 'examination' were computed in an analysis with the teachers corresponding to each of the six tests, and the students, as actors.

The values are very stable over the years. The somewhat lower values of the very last test might be explained from the possibility of compensation: students having good marks on other tests might risk to get only a '5' on this one).
However, the interpretation of differences between values of marks for different tests or different years is not the most interesting use that can be made of this particular model. The computation of values is only the starting point of the analyses. The point is that the model makes it possible to construct a conceptually very different set of new variables from the known set of data on time spent and marks received. The new variables are individual characteristics, like the interests of individual students in marks and in time, and characteristics of the system, like the values, but also the particular distribution of interests over students. Also the model differs from traditional methods of data analysis, because it is now possible to generate *predictions* of the behaviour of the *system*, and of the individuals in the system, when there are particular changes in available budgets, or when the group of students is selected on particular characteristics. In this way it might be possible to characterize the sub-group of drop outs as differing in certain respects from the group represented in Table 1. Of particular concern is the role of the test as a clearing house for the exchange of marks and time: clearly test results must be predictable for the individual student, otherwise investing time in preparation for the test can not be seen as resulting in higher expected marks. Many interesting analyses remain to be done, some of them to be reported on in the paper. Also there are additional data available on the marks expected by the student, and on time spent in class and in group work.
What will be highlighted in the paper is the competitive character of this educational system. It is not just the teachers who are competing with each other for the time of the students, but also the students are competing with each other for the scarce resource called 'marks'. Drop out rates and grade retention rates, as characteristics of the educational systems involved, depend in complex ways on the behaviour of individual actors. Social system theory, in contrast to traditional linear models for data analysis, allows individual actors to compete with each other, it does not assume that individual actors are independent of each other. This theory might be used to unravel the perennial problem in educational research: the relation between individual actions at the micro-level, and educational outcomes at aggregate levels. Educational assessment is a prime example: lacking absolute standards, assessment is done on a relative basis. Where teachers use relative norms, educational innovations cannot possibly result in better examination results. The reverse may also be true: teachers using relative norms in the grading of students may loose control over the situation, resulting in study delay and attrition.
The application of Coleman's social system theory to examinations, or the combination of tests comprising the particular examination, results in a meaningful *description* of this educational system. Such a description might be an excellent basis for the application of more *prescriptive* models, such as Van Naerssen's (1974) model that prescribes what optimal strategic behaviour of students is, given a particular testing situation.
The work presented will be discussed in several ways. First, does this kind of data analysis result in a meaningful connection of the micro-level and the system-level (what the theme-coordinators call the meso-level)? Second, what further research

hypotheses does the presented analysis generate? (for example: concerning the use of the Van Naerssen prescriptive model). Third, what policy advice is immediately evident from the analysis as presented?
In the discussion the results of the social systems analysis will be contrasted with earlier advice on examination policy, given to the department of law, and that has stimulated discussion within the University of Amsterdam about examination procedures. Some remarks will be made about the 'Self-evaluation project', the project that generated the data reported on above, because it is related to the actual debate on the possibilities of advising particular students to drop their study, given poor test results during their first year. The project was meant to give the students some insight into their academic situation at the end of their first year of study, using data they themselves provided on marks expected and time spent, and related to marks received (see Voorthuis & Wilbrink, 1987, chapter 4, for a summary report on the project).

References

Coleman, J. (1990). Foundations of social theory. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Naerssen, R. F. van (1974). A mathematical model for the optimal use of criterion referenced tests. *Nederlands* Tijdschrift voor de Psychologie, 29, 431-445.

Voorthuis, M., & B. Wilbrink, (1987). Studielast, rendement en functies propedeuse. Relatie tussen wetgeving, theorie en empirie. Deelrapport 2: Evaluatie-onderzoek Wet Twee-fasenstructuur. Amsterdam: SCO.

SELECTION OF STUDENTS DURING FIRST YEAR OF HIGHER EDUCATION

Z.C. Huijsman, C.H.J. van Veen, & H.K. Letiche, The Rotterdam School of Management, Erasmus University, Rotterdam, The Netherlands

Summary

In the Netherlands, the first year of higher education, which is called the "propedeuse", determines which students will continue further in their education. This decision is based on their academic performance. The article describes an investigation into the criteria of academic success in the propedeuse of the Rotterdam School of Management, Erasmus University, and discusses alternative scenarios for the propedeuse. Because the study is still going on, we are only able to describe in this summary the goals of this study, its design, and some preliminary results. At the conference we will present some of the final results together with a discussion of their significance.

The need for a descriptive study stemmed from a felt lack of information about the effectiveness of the propedeuse. The Management School has experienced a relative high level of student dropout late in the program. An effective propedeuse should select students to continue their study who will demonstrate a high level of academic, skill–based and attitudinal competence, leading to the granting of the doctoral degree (MSc). The overall criteria for selection must be clearly grounded in goals specific to the program. Because the focus of this study is on selection of students in the first year, causes of dropout later in the program will not be dealt with specifically.

The criteria of academic success are described by viewing the curriculum of the propedeuse from the different forms a curriculum can take (derived from Goodlad, 1979): *formal*, the official documents describing the curriculum; *perceived*, what lecturers think of it; *operational*, what lecturers organize and do in order to create a learning experience for students and the role of the exam system and regulations; and *experienced*, the learning experience of students. To describe these different forms of the management curriculum, different methods of data collection were needed:

- analysis of documents, study materials, student work, and the exam system and regulations;
- observation of lecturing;
- open–ended interviews with faculty, pedagogic support staff and first year students.

1195

The descriptions of the different forms of the curriculum show that the learning experience of students was not always the same as what was planned in the curriculum documents. Lecturers are autonomous in the way in which they design the learning experience for their students and in how they measure outcomes. Most lecturers make use of multiple choice exams to achieve efficiency. Lectures often don't know what their colleagues are lecturing. Ideas about management education and how students should study management differ widely among staff. Not surprisingly, students complain about overlap in course content and the diversity of standards by which their academic performance is measured. Nevertheless, most students report enjoying the first year because of the diversity in course content and teaching methods.

The decision whether a student graduates to the second year, is based on whether the student passes his/her courses (i.e. on exam results). During the last three years the requirements to pass (the exam regulations) have been gradually stiffened, to an extent that, at the moment, a student has to pass all courses and no compensation between course grades is possible. This means that courses where traditionally many students fail, such as quantitative methods (mathematics and statistics), have become even more an obstacle, because students can no longer compensate their grades in these courses with grades in courses where almost all students pass.

Because most courses consist of knowledge transfer via lecturing and reading, and measure student performance with multiple choice exams; the criteria of academic success are closely related to the ability to digest theoretical knowledge and to reproduce it when asked to. The practical orientation where in students are asked to use their knowledge to discuss issues and to analyze problems in order to find practical business solutions is present, but because it is difficult to assess skills required, these course do not form a determinant of academic success in the propedeuse.

The process of criteria formulation for the freshman year could, we believe, better follow the next steps:

- formulate the profile of the graduate;
- translate this profile into a development path which describes knowledge, skills and attitude development;
- formulate what basic knowledge, skills and attitudes should be present when students start their studies.

Ideally, the criteria of academic success ought to combine the ability to acquire theoretical knowledge from several disciplines, with the use of this knowledge in discussing issues and analyzing problems. Therefore proposals for change will need to focus on the way management learning takes place, whereby students ought to be challenged to gather knowledge leading to an understanding of key features in the business world and to analyze and solve managerial problems.

References

Goodlad, J.I. et al. (1979). Curriculum inquiry: the study of curriculum practice. New York: McGraw–Hill Book company.

DEVELOPMENT AND TESTING OF A CAUSAL MODEL FOR ASSESSING PROBLEMS OF DUTCH RESEARCH ASSISTANTS (PHD STUDENTS)

J.F.M.J. van Hout, & H.J.H. Jurgens, Centre for Research and Development into Higher Education (IOWO), University of Nijmegen, Nijmegen, The Netherlands

Introduction

The researches reported on in this paper deals with the problems of postgraduate research assistants in the Netherlands. A first survey was conducted in 1983 before the implementation of a new educational system for obtaining a doctorate, equivalent to a Ph.D., the so-called 'assistants-in-training system' (AiO-system). In a second survey, conducted in 1990, this AiO-system was evaluated. In this paper, the final causal model obtained from the first survey will be tested on the data of this second survey.

Theoretical framework

On the basis of the findings in the review of the literature and the limited Dutch literature available on research assistants, a categorisation of possible *problems* experienced by research assistants is developed:

- Research problems with respect to (1) choice of topic, (2) time planning, (3) quality of the research, (4) facilities, (5) writing the thesis.
- Teaching problems with respect to (1) presentation skills, (2) structuring of the subject matter, (3) motivating and activating students, (4) sufficient mastery of the subject matter.
- Problems in the working environment with respect to (1) supervision, (2) the relationship with the department, faculty and students, (3) the position and status of the research assistant, (4) the formal legal status.

The literature does not provide a clear insight, either theoretically or empirically, into the relations between causes and problems. Because of the large number of variables many relations would in fact have to be interpreted as quasi-relations. Therefore, the variables should be tested in connection with each other. In view of this we have constructed a preliminary global causal model for problems experienced by research assistants. The basic assumption was that the model has to be as simple as possible.

1197

The first survey

In 1983 166 research assistants in three faculties - natural sciences, social sciences and humanities - of six Dutch universities filled in a questionnaire (response 66%). This survey has yielded the following variables via factor and scalogram analyses.

- Individual characteristics: age; sex; teaching experience; years of service as research assistant; teaching motive; university research motive; teaching orientation; and research orientation.
- Characteristics of the working environment: faculty (natural sciences, social sciences and humanities), teaching orientation of the department; research orientation of the department; supervision; planning; working autonomously or in isolation; and extent of the teaching assignment.
- Problems experienced: teaching concerns; teaching stress; research concerns, research stress; satisfaction with the supervisor; satisfaction with the department; and satisfaction with the status of research assistant.

The preliminary global causal model has been operationalised and made concrete by means of a number of hypotheses with regard to the relations between the variables. The operationalised model was tested via the LISREL-VI procedure: a procedure designed to test causal models. This procedure showed that the model had to be rejected since it did not fit the relations as found between the variables. On the basis of the relations observed in the literature and the parameters of the LISREL procedure it was possible to improve the model in incremental steps. The model that was finally obtained rejects the basic assumption on which the preliminary model was constructed, but not its specific hypotheses. The improvement consists mainly in the adding of relations. The final model is highly acceptable, both with regard to content and from a statistical point of view (chi-square 152 with 165 degrees of freedom, p=.74, with a 'goodness of fit index' of .929).

The second survey

In a survey, held in 1990, Ph.D students (n=877) and their supervisors (n=550) were interviewed. The purpose of the research was, on the one hand, an evaluation of the formal implementation of the AiO-system and, on the other, a description of the training, counselling and research progress of the Ph.D student. A large number of questions were phrased equivalently to the ones in the 1983 survey. This offers the opportunity of testing the final model obtained from this first research.

The results of these analyses will be presented at the conference.

References

ABRC (Advisory Board for the Research Councils). (1982). Report of the working party on postgraduate education. London: Her Majesty's Stationery Office.

Berelson, B. (1960). Graduate education in the United States. New York: McGraw-Hill.

Bucher, R., & Stelling, J.G. (1977). Becoming professional. Beverly Hills: Sage Publications.

Cartter, A.M. (1976). Ph.D.'s and the academic labour market (The Carnegie Commission on Higher Education). New York: McGraw-Hill.

Hay, A. (1987). The contribution of UK postgraduate thesis research to the published literature of four social sciences, 1980-1982. London: ESRC.

Katz, J., & Hartnett, R.T. (Eds.). (1976). Scholars in the making: The development of graduate and professional students. Cambridge: Ballinger.

Malaney, G.D. (1988). Graduate education as an area of research in the field of higher education. In J.C. Smart (Ed.), Higher education: Handbook of theory and research (Vol.IV, pp.397-454). New York: Agathon.

OECD (1987). Post-graduate education in the 1980s. Paris: Organisation for Economic Co-operation and Development.

Rudd, E. (1975). The highest education: A study of graduate education in Britain. London: Routledge & Kegan Paul.

Rudd, E. (1985). A new look at postgraduate failure. Guildford: SRHE & NFER-Nelson.

Solmon, L.C., Kent, L., Ochsner, N.L., & Hurwicz, M.L. (1981). Underemployed PhD.'s. Lexington Mass.: Lexington.

Vartuli, S. (Ed.). (1982). The Ph.D. experience: A woman's point of view. New York: Praeger.

Welsh, J.M. (1979). The first year of postgraduate research study (SHRE Monograph 37). Guilford: The Society for Research into Higher Education.

Winfield, G. (1987). The social science PhD: The ESRC inquiry in submission rates (The Report). London: Economic and Social Research Council.

A USABLE APPROACH FOR EVALUATION IN HIGHER EDUCATION

Gerard van den Berg, Educational Development Unit, Delft University of Technology, Delft, The Netherlands

Introduction

In the last few years Higher Education has shown an increasing interest in evaluation, partly due to the coming of review committees. Unfortunately in practice evaluation is (all) too often reduced to gathering information. It so happens that implementing educational innovations and/or changes on the basis of evaluation data turns out to be a laborious process. Especially if the initiative for evaluation is taken on a higher level (faculty board), it is often the lecturers who frustrate an actual change of educational practice.

Some possible explanations for these experiences can be put forward from theory. The characterization of a university as a professional bureaucracy with departments as autonomous units (within which lecturers, in turn, are often autonomous components) has to be taken into consideration. In addition a model which shows how lecturers might deal with data from student evaluation of university teaching can be of help.

On the basis of these theoretical considerations an alternative approach for programme evaluation is proposed. On the one hand this approach seems to be feasible and on the other hand it might provide a suitable basis for the implementation of actual changes in educational practice. The components of this approach are not new (of course), but the way in which they are used is. These components are: information from our student data system, (standard) questionnaires, lecture response groups, examination results and a subsequent meeting in which all of these data are discussed with students, lecturers and department together.

Design and Implementation

The new approach has been tested in the faculty of electrical engineering for their first year-programme. The basis for the programme evaluation is the subsequent evaluation discussion for each subject. These meetings are attended by the lecture response group of five students, the teaching team, a member of the educational bureau of the faculty and a faculty administrator. The discussions are directed by an outside educational specialist and are conducted on the basis of a list of items that need attention. The students participating in these discussions are considered to be the representatives of the whole studentpopulation. Their

opinions about the various points of attention should not be personal but should be tested by the opinions of other students.
In consulting the lecturers about this procedure some resistance was sensed. Some of the lecturers did not wish to be pinned down by the opinions of only a few students. A quantitative backing up was found a necessity. To take away the resistance and to create a sound basis for the evaluation in the faculty a standardized student questionnaire was proposed.
This questionnaire should be considered as a signalling system to be used in the evaluation discussions. Both students and teaching team should give their explanations about and/or comments on remarkable results. The questionnaire was drawn up from a questionnaire that was already in use in the faculty. The lecturers were allowed to comment on and to supplement this questionnaire. All items were of Likert-type with a 5-point scale. For most of the items it was not possible to formulate a standard beforehand. The evaluation discussion, then, should also serve the purpose of setting a standard for the future. The students' union was called in to inform students about method and execution of the evaluation.

Some results
One of the most remarkable things we found was that when the lecturers' mistrust about the purpose of the evaluation approach was taken away, their strong responsibility for the quality of education became obvious. For example, they were very approachable for complaints about the high pace and overload of their subject. Lecturers did promise to use this information in the revision of their subject for the next year. They kept their promise and a better matching of student workload and creditpoints was the result.
The evaluation discussions also resulted in the publication of a student learningguide for each subject. At the same time the faculty administrator promised that the preconditions for teachers would be improved. As a result groupsize was better adjusted to teaching methods and more suitable instructional rooms became available. So a series of practical agreements were the benefit of the evaluation.

Some concluding remarks
It is obvious that, after one try-out, a conclusion about the functioning of this evaluation approach must be tentative by nature. The proposed evaluation approach has to be considered a qualitative system in spite of the quantitative data from the questionnaire. Although it is not possible to give quantitative effects of the new approach, we have seen actual changes in education as a result of the evaluation discussions. The evaluation discussions themselves are an instrument for faculty development. They can be typified as an effective way of professionalizing all the people concerned in education. What is effective in particular is that a situation has been created in which students, lecturers and administrators can make their demands/wishes subject of a mutual discussion right away. Since making issues a subject of discussion has led to actual changes in educational practice, this approach seems very usable to me.

1201

References

Berg, G. van den (1990). De evaluatie van het eerste jaars onderwijs op elektrotechniek. In: Jochems, W. (red.), Onderwijsverbetering, voorbeelden uit het wetenschappelijk onderwijs. Delft: Delftse Universitaire Pers, 69-78.

Mintzberg, H. (1988). The professional bureaucracy. In The strategy process. Quinn J.B., Mintzberg H., James R.M., Prentice-Hall.

Neve, H.M.F. de (1991). University teachers' thinking about lecturing: student evaluation of lecturing as an improvement perspective for the lecturer, Higher Education, 22, 63-91.

1202

VIEWS ON 'PEER REVIEW' AND QUALITY ASSURANCE

Janet Powney, Scottish Council for Research in Education, United Kingdom

Introduction

The use of peer review as a tool of quality assurance is familiar throughout education - in teaching, research, course validation and accreditation. This paper draws on detailed empirical studies in U.K. higher education concentrating on characteristics of peer review emerging from course validation. These characteristics undoubtedly share common features with other uses of peer review.

Who are the peers? How do they assess quality in valid and credible ways? Can peer review survive as a useful activity in a society which increasingly expects higher education system to be typified by competition rather than co-operation? These are some of the questions which will be underpining the discussion.

The predominant focus will be on peer review as developed in the UK from 1965 by the Council for National Academic Awards (CNAA), a system which is to be disbanded following recent government legislation. The CNAA experiment was" *a particularly British system of peer review of courses"* (Church 1988). Nevertheless it is reasonable to hazard that this system has implications for the use of peer review in all sectors of higher education throughout Europe since it has been designed to engender and enhance not just professional responsibility but also that very 1990's feature, accountability, among academic staff.

Between 1964 and 1992, the model of peer review in the public sector (i.e. the non-university) sector of higher education required these polytechnics and colleges of higher education to draw substantially on external referents for quality, peers in universities and among practitioners in occupations and professions. This model not only gained the confidence of those working within the colleges and polytechnics, it also gained respect, and is to some extent being emulated by those who in the past have been its judges, the universities. The CNAA experiment can be characterised as peer review built upon *openness, power sharing and academic and professional responsibility* all of which were intended to enable institutions to develop academic autonomy arising from collegial responsibility. The system could also (in theory at least) *empower* its participants to take charge of their institutional direction.

Peer review makes an important contribution to the quality of an institution - and by quality I am using the "fitness for purpose" concept (Ball 1985) and leaving aside debates about what should be the purpose of higher education. Whether

1203

scholarly, vocational or directed towards holistic views of educating the 'whole person', quality must be concerned with that which is intended. Since peers are by definition those who share the same interests, knowledge, experience and/or the practice of the subject under consideration let us assume they share similar purposes. A scholarly programme will invite the opinions of other scholars on the knowledge, teaching and research with which supports its students. A vocational course may prefer the opinions of practitioners in the field as being more appropriate for the task.

Peer review is used in different forms in the USA, Europe, Australia and other countries to deal with validation of individual courses and programmes as well as with accreditation of whole institutions. Martine Neave (1991) provides instances of how quality, the shared concern or western European countries is in The Netherlands, Belgium and France being verified by peer review in various forms.The imperative is for the various academic awards to be maintained at standards which ensure comparability (and therefore currency) with other awards at the same level, i.e. that doctorates represent a level of academic achievement common at least throughout a country, possibly across international boundaries. It is of course impossible to reach exact comparability but at least there is consensus on the acceptable range of attainment for an award. The level of an award such as a doctorate also has to be distinct from other awards, eg diplomas.

For a variety of reasons, there can be no absolute guarantees about the quality or level of an award but peer review should ensure that all the inputs to a course from the institution are sufficient for students to be likely to succeed. An institution should have confidence that sufficient resources in terms of qualified staff and facilities have been provided for all its validated courses

UK Background

When higher education expanded rapidly in the 1960s and 1970s in the UK, along with the principles of more open access to students, as expressed in the Robbins report, there was an urgent need to increase the number of higher education institutions. Rather than merely expand the size of universities, the government set up new vocationally oriented institutions, polytechnics (DES Robbins Report 1963). Over their brief 25 years' existence, institutions have maintained a vocational element but there has also been 'academic drift' towards the kinds of courses already provided by universities, partly because many perceived the polytechnics as being of lower status than the universities — the opposite position is I believe true in France. Recent legislation has not only abandoned the deployment of CNAA. It has also enabled polytechnics to change their titles to include the term 'university'.

What has been the process which enabled the polytechnics, and some other institutes of higher education, most of whom had no previous experience of higher education, to reach university status within some twenty five years? (I shall leave aside the issue of whether or not this was an appropriate ambition.). At the root of this development is a system of peer review conceived as a means of collective maintenance of standards in institutes with teaching rather than research as their main concern. It was initially dominated by those outside the institution and was gradually taken over by peers from inside the organisation as the institution became self-validating (see Adelman & Alexander 1982). At first individual

courses were validated and when a polytechnic or collect of higher education had built up a portfolio of validated courses, the whole organisation could apply to become an accredited institution of CNAA in the same way that an American college would be accredited by its State.

CNAA validation

For those not familiar with the validation system, a team of academics put together a proposal for a course delineating its aims, objectives and an outline of content as well as details of how it will be taught, resourced and assessed. This proposal is scrutinised in its written form by a panel of peers drawn together for the sole purpose of validation of this programme. Panel members discuss the proposals among themselves and then meet with the course team and the management of the supporting institution to explore further details and to decide whether or not the students studying on the proposed course are likely to meet the standard of the award — certificate, diploma, degree, masters, etc..

Periodic reviews of validated courses usually take place every four or five years and follow a broadly similar pattern with a specially convened panel of peers who review evidence on the nature and performance of the course. Panel members do not observe the course in action — they neither, for example, attend lectures or seminars, nor do they attend assessment boards. Course teams are expected to provide performance indicators and other evidence including reports from external examiners* on how effectively the course runs and once again the panel of peers has discussions with the course team. At this stage there is also the benefit of comment from existing and past students of the course. Overall, the validation and review procedures have been seen as an important aspect of enhancing quality and improvements in the design and content of higher education courses.

Characteristics of peer review

The polytechnics where I have carried out my empirical studies all followed the principles on which CNAA based their practices and which I have identified the key characteristics as: openness, power sharing and professional and academic responsibility. These characteristics are essentially concerned with empowering participants and have student experiences as their anchor.

There are almost invariably gaps between intended practice and what actually happens. Close examination over more than five years of detailed case studies of review and validation events in polytechnics, as well as other empirical observations and interviews of a wide range of participants within the system and outside observers, provide some evidence of how the participating of colleges have been empowered; there is also evidence of conflict between bureaucratic demands and the provision of intellectual debate and innovation.

---

* Every academic course leading to a UK academic award appoints one or more external examiners from other institutions to ensure that assessment procedures are carried out correctly and that the overall level of assessment of student work is at a comparable level with that of other institutions offering the same award.

1205

Openness

The first essential ingredient is openness, transparency of the process and high visibility of all the participants. Proponents of courses build their courses on explicit criteria conforming to the spirit of the guide-lines provided by the Council for National Academic Awards (CNAA). Peer review "uses the politics of mutual education so that conflicts of value are open to critical debate". (Simons 1989) Billing (1986) suggested that this was a "process of legitimation ... a way of building consensus internally and externally by adding together a number of subjective judgements to develop confidence in degree courses among tutors and outsiders".

The level and nature of this open debate will depend on who takes part and it is evident that tutors most value the contributions of subject specialists rather than other kinds of peers who make suggestions and advise on topics of general concern (such as course organisation and control) without being subject specific. Openness is more likely to occur when there are shared and not competing interests. In the 60s and 70s in the early phases of CNAA, there was no shortage of students and peers from other institutions who were generous in sharing their expertise freely. As resources available to higher education have declined so tutors have had less time to spend on broader professional activities outside their teaching, administrative and research tasks. Fewer resources distributed in ways that encourage competition between, and ranking of, institutions who previously collaborated, has shifted priorities. The emphasis now is less on collegial responsibility and more on personal career enhancement which fits in with schemes of appraisal of the individual and assessment of the institution for distribution of resources for research grants.

Power sharing

Power sharing had several dimensions. Courses that were approved had to be proposed by a team of people not by a single academic and the validation process was geared to judge whether or not the course team was sufficiently together to produce a coherent programme where progression from one part of the programme to another made good logical sense. (This drive for logical coherence broke down with the growth of modular schemes with free standing units.) The team, therefore, shared responsibility very explicitly for the quality of the course.

Students as 'informants' and 'judges'

A further interesting dimension of power sharing has been the high profile given to students' views initially on the suitability and then on the delivery of a course. CNAA always required some contact and discussions with students, not as part of the learning/teaching process but discussion on the learning/teaching process. Consistent with the notion of power sharing and student involvement has been the invitation to students to become members of peer review panels. Lancashire Polytechnic espoused the principle of including students on review committees in the late 1980s (Coyle and Powney 1990) and, although there were some logistical difficulties, continued to promote the contribution of students "as members of the polytechnic and therefore also responsible for the quality of the education provided".

To what extent, however, are all roles in the validation and review process open? When 'peers' = 'subject specialist' there is less emphasis on organisational status and more on academic expertise. As institutions have taken on responsibilities of self accreditation, panels have been dominated by internal members who can contribute to general issues such as resource questions and management issues. They are necessarily experienced and senior members of the institution. In UK higher education, these are predominantly and disproportionately white males. They are in "the great and the good". Only when a programme being validated concerns women's or multi-cultural issues is there a less Caucasian male dominance.

Conclusions

Peer review has potential problems based on:

- the sub-text or alternative agenda may be more important than the explicit agenda and record of events;
- the choice of peers may be narrow and inclined to promote the status quo rather than innovation;
- unnecessarily bureaucratic procedures which distract attention from the key point of quality, i.e. the delivery of the course;
- being predicated upon generous collaboration/co-operation.

It is less likely to flourish in an environment in which institutions which are part of the review system are competing a) for resources and b) for students.

In spite of these caveats, CNAA peer review has proved useful in establishing the credibility of public sector institutions in the UK, thus fulfilling its major task. As a by-product it has also promoted rigourous approaches to monitoring and review to benefit recently accredited institutions (the ex-polytechnics) and long-established autonomous universities.

References

Adelman, C., & Alexander,R. (1982). The Self-Evaluating Institution:practice and principle in the management of educational change. Methuen.

Adelman, C., & Powney, J. (1987). Institutional Agreements and the Validation Process. Evaluation of the one year Agreements between CNAA and Newcastle- and Sheffield Polytechnics. CNAA.

Ball, C. (1985). Fitness for Purpose. SRHE/Nelson.

Church, C. (1988) 'The Qualities of Validation' in Studies in Higher Education , 13, No.1.

Coyle, P., & Powney,J. (1990). "Informers and judges - students in course review and validation" Paper presented to SCED Conference, Newcastle Polytechnic.

Neave, M. (1991). Models of Quality Assurance in Europe. Council for National Academic Awards, London.

Powney, J. (1988). Rapporteurs at Portsmouth Polytechnic. First Report of CNAA Funded Project on Performance Indicators. Portsmouth Polytechnic.

Williams, G. (1990). "Quality and Resources Allocation" in Loder,C. Quality Assurance and Accountability in Higher Education Kogan Page.

THE EDUCATIONAL RESEARCH AT THE NATIONAL UNIVERSITY OF MEXICO

M.T. Poschner, Educational Research Centre, Centro de Investigaciones y Servicios (CISE), Unam, Mexico

Introduction

This research was conducted in the National University of Mexico after observing almost a total absence of any study about the state of art of its educational research. At the beginning of the research, the purpose was to acquire more detailed information about it and therefore we started an inventory in all the schools were we could find academic personnel pursuing any research project. Although the period in which this work was done was between 1983-84, we consider that its results have validity also nowadays because it is not yet very usual that the teachers make educational research together with their teaching praxis. We have considered as "educational research" all types of research based on the study, description or analysis of the educational reality either in a multidisciplinary or an interdisciplinary way. Our main hypothesis was that it is through the development of educational research by the teachers themselves that it is possible to improve and ameliorate the quality of teaching and the education in general in our University. On which theoretical position do we ground this asseveration? To explain this we need to get back to the beginnings of our educational research practice which was within the field of the inservice teacher education. We observed (together with another colleague) that it was not enough to give the classical "training" courses to the teachers so as to bring about significant changes in their teaching by observing it and doing an inquiry work over it to improve their teaching -which is also ours- can be found in the works of Lawrence Stenhouse, De Lansheere, Carr and Kemnis, Huswn and others. We built up our "ad-hoc" action research method on the experiences we had in our courses with the teachers and on the theoretical contributions of Paulo Freire, Anton De Schutter, Orlando Fals Borda, Heinz Moser, and much more people who made valuable field work in Latin America and other countries of the Third World.

Method

We have begun this research with a general and deep inventory of all the research made by teachers as well as by researchers. There were very few records about these activities in the schools and centres and therefore we have often received information in "chain" form (snowball): from one school we "discovered" the other ones. Each teacher/researcher who had at least one project

was interviewed with several questions about the institution and its support to the educational research. The school authorities we could find were interviewed with several questions about their information level and their attitude towards educational research. The quantitative analysis of the data of the research projects was made on the behalf of basic statistical methods. The qualitative study of the answers of the teachers/researchers about their research projects was made to obtain afterwards a classification out of the different types of research and the focus of their application in the educational context. It was also made a "content analysis" out of the questions about the relation between educational research/authorities/decision making/educational policy in the university and about the different problems on making educational research parallel to their teaching praxis inside the University.

Results

We detected in the National University of Mexico (the biggest one in the county and mostly part of the world based on the number of students) 33 schools and institutes where teachers and researchers made educational research. The total number of projects was 582 developed by approx. 470 teachers/researchers. The topics of their research works were about education in general (10%), about the University itself (14%), about the curriculum (41%) and about the teacher-pupil relationship and social context (35%). The research types were descriptive (32%), evaluative (45%) and propositive (23%). We found that the interest of the teachers/researchers was focused mostly towards the evaluation of the curriculum and the teacher/pupil relationship and the social context. These are the main results but the study gave a lot of more information which we do not mention here now because of its great number.

Conclusion

We observed that there is strong interest in the vast majority of the teachers/researchers in Mexico in general (based on other studies) and also in our University in particular, to investigate their own educational context and to search for possible solutions to the different problems in the teaching and learning process they have detected through their research works. This way of working in the teaching process from the inside of the University working in the teaching process from the inside of the University is an indicator of the fact that the educational research is conducted not only by researchers but also by interested and active teachers who want to bring about changes in their own praxis and their educational context. We can see a gradual demitification process in the educational research area and therefore that it is needed a vast work and special strategies on the preparation and development of teachers on educational research methodology. It is important that the academic personnel can get conscient of the value of the educational research made by themselves so as to get out of the micro world of the classroom and arrive to a macro level analysis on the possibility that the institution's authorities benefit and take profit from the results of the research works to the decision making in its educational policy.

1209

Discussion

We consider that it is through these kind of studies that we can learn what is done inside the university, which are its problems and above all how to find the possible solutions by offering better procedures or presenting innovative applications on the teaching & learning process directly mentioned by its academic personnel. The possibility to influence the university's authorities on using the research work results to decision making and to build educational policy in the future is a problem which was very often discussed all over the world by higher educational experts and we haven't got yet a solution because it is a very complex question in which we must consider different aspects of the social, economical and political context of the countries in which we are working,.In this sense our position is that we must continue working together with the academic personnel on the different matters of the educational research (e.g. organizing internal and external networks, special courses, round tables, seminaries, special discussion,groups, etc.) so as to help improvement on the quality of education research in general and of the teaching-learning process inside the University and of the future professionals in the county. Finally we can see that this humble research begun with a simple inventory researched to be a quantitative & qualitative research with the intention to help the academic personnel in the decision making process about the detected educational problems so as to improve the quality of education in general and to better education in Mexico itself.

FINANCING OF STUDENT STATUS: ANGLO-GERMAN COMPARISONS

Alan Brown, University of Surrey, United Kingdom; Martina Behrens, University of Bielefeld, Germany

Introduction

This study was part of a much broader research programme which looked at the transitions to adulthood of 640 young people in two pairs of matched towns: Liverpool and Bremen; Swindon and Paderborn. The longitudinal study followed these young people for between three and five years. 40 young people in each town were initially identified as being on academic routes likely to qualify them to enter higher education (HE). Others were on vocational routes which could, and sometimes did, lead to entry into HE. From all such potential entrants to HE we collected information on their career plans and whether they intended to enter HE at some stage. An issue frequently mentioned was finance. As a consequence we pursued the theme of how young people finance their studies in HE.

Method

We already had detailed survey and contextual information about the lives of our sample from age 16 or 18, including experiences, attitudes and aspirations towards education, training and employment. After following their progress for several years we carried out another survey in 1991 of a subset of 40 young people (then aged around 22) in each town. Altogether about a third of that sample had either attained HE entry qualifications or else still harboured aspirations to enter HE. Our quantitative and contextual information was then supplemented with ethnographic interviewing of 16 young people in each town, with about 25 having either entered HE or given it fairly serious consideration. Such information then allows us to give a fairly thick description about the processes whereby young people choose to enter HE and the role of finance in that decision. The careful matching of the wider sample by age, gender, initial career trajectory and labour market provides an illuminated backdrop against which to consider the financing of HE study for this age cohort in the four towns.

Results

Finance as a consideration in going into HE

For some students in all four towns the decision to enter HE was unproblematic. Their own aspirations and the expectations of school, peers and parents meant that it was seen as a natural progression, and how to finance their studies did not

1211

loom large in such decisions. For other prospective students, however, finance was a key consideration. In the most extreme cases, lack of sufficient financial support and the need to contribute to family income were seen as effectively ruling out the possibility of entering higher education for the foreseeable future. In all four towns there were examples of some young people making assessments of whether they could afford to go into HE. Also in England, some of those for whom finance was not an issue in the initial decision to enter HE, became increasingly concerned about money as student grants were sharply reduced in real terms over the last five years.

Financial support of young people in HE

The German system of higher education, in which no fees are charged and grants seldom awarded, relies partly upon a system of student loans. However, parental support is much more significant, followed by part-time employment, with only a minority of students receiving any state support. In England, grants, banks and parents all figured as significant forms of external support. There were quite wide variations as to how much of students' income came from grants. Parents and banks were the most popular external means of supplementing the shortfall caused by what the students in England almost universally saw as the inadequate level of grants. Under the grant system, parents had always been expected to make up the grant with a parental contribution, the size of which depended upon their income. For a long time, some parental contributions had fallen below the expected level, but the problem was exacerbated by the falling level of grants in real terms.

Working to finance HE study

Students may, of course, finance their own studies, either in part or in whole. Working was one possibility, although most students who were successful on the academic track did not actively consider going to work before entering HE. Whether prospective students do actually go through with their intention of going into HE after working depends both on the nature of the work and their commitment to a chosen career path which involves HE. Thus two young people from Paderborn entered work with the intention of going into HE later found the attractiveness of work fatally weakened their resolve. There was clearly a labour market effect in that the existence of relatively well—paid jobs and apprenticeships, particularly in banking and insurance in the more buoyant labour markets of Swindon and Paderborn, meant that it would be very much easier for prospective students to get such work than those from Bremen and Liverpool. While some students worked prior to their studies, very many more financed their studies through working during vacations and/or part-time while studying. In England, the relatively speedy student transitions did mean that for undergraduates at least, it was relatively rare for part—time work to slow progress towards a degree. In contrast, lengthy transitions and periods of work often went together for German students.

Discussion

Overall in both countries parental financial support is becoming increasingly critical to the financing of student status. Those with significant parental financial support can focus upon other aspects of managing student status, but for those without such support financial issues can become the dominant concern of that management process. This is particularly the case where the need to work has to be balanced with the requirements of study, but it can also operate as a pervasive influence affecting most aspects of student life. The attrition of those with the capability but not the resources to benefit from HE at both before and during their student career should be a major public policy concern.

The extended transitions of German students can become a test of endurance for those without substantive parental support. The lesson for Germany would then appear to be if clear bounded signals are given about achievement in specified time—frames this might itself act to speed student progression and transitions. The speed of English transitions, typically three years, means that for those who have decided they can afford to enter HE then there is a clear end—point in view. Students may accumulate considerable debts along the way, but graduation remains a fixed point for most students.

More radically perhaps, both countries could reassess their expectations of 'full—time' students. That is, in England institutions could recognise that significant numbers of students will have to work contemporaneously to finance their studies and consider the implications this could have for all aspects of the curriculum they offer. Whereas in Germany, the converse would hold: rather than the open—ended injunction to take the final qualification when you are ready, a clearer negotiated timetable which recognises achievement in medium rather than the long term would be desirable. Whether these are attempted, it should be recognised that financial issues need to be considered from the outset when formulating or reviewing policies towards HE access and progression. Furthermore, they need to be considered in the context of education and training policy for young adults as a whole, not as a separate sub—set of issues.

1213

THE PLACE OF PLANNING IN LEADERSHIP EDUCATION

Antonia D'Onofrio, & Marian Fox, Widener University, Chester Pennsylvania, USA

Nature of the Problem

A national survey of professors of higher education leadership demonstrated how planning is taught in post-secondary administration programs; and whether current teaching methods integrate strategic and tactical concepts and skills (at the meso level of planning).

Perspectives from the Literature

Planning theorists frequently distinguish between the strategic tasks of setting broad goals and narrower tactical objectives, (Cope, 1981 and 1987; Miller, 1983; Morrison, Renfro, & Boucher, 1988). These researchers assert that important planning processes might include the negotiation of stakeholder concerns, galvanizing the will to change; developing long term perspectives on the structure and operations of academic programs; recognizing political, social, economic and cultural contexts that shape change; and focusing resources on new goals without incurring additional economic or reputational risk.

The Conceptual Model

When management is tactically engaged, it fosters interactions between individuals, individual departments, and linkages between groups. Its defining activities include limiting roles, centring on specific tasks and duties, limiting risk, preventing errors, conserving resources, lowering uncertainty, and reaching closure on decisions. when management is strategically directed toward broad external influences, it focuses on gathering information needed to bring group tasks to completion, realizing common goals with other organizations, and profiting from shared ventures. Its defining activities include broadening responsibilities, tolerating risk, exploiting resources, and adopting a hypothetical approach toward decision-making.

Methods

Higher education faculty were asked to describe the planning concepts, methods and instructional resources they emphasized in leadership courses on a three-part survey. The first part of the survey consisted of 24 forced-choice triads designed to distinguish faculty who emphasized internal stresses and adaptations from those who stressed broader influences on organizational plans. Triads consisted of three

statements posed as instructional objectives. Faculty had to rate each objective, indicating which objective was Emphasized Most, Emphasized Least, or Neutral. Forced-choice options were used to control most response biases associated with rating scales. The second part of the survey required faculty to indicate the kinds of planning methodologies they favoured from a list of 17 commonly employed methods. Methods ranged from non-quantitative approaches (eg. nominal group techniques) to highly quantitative approaches (eg., trend-analysis). The third part was, open-ended; respondents listed major authorities and texts that influenced their teaching of planning.
Magnitude scaling was used to discriminate disparities of opinion about instructional objectives that were either stressed or discounted when teaching planning. Favourite planning methodologies were enumerated. Open-ended responses were content-analyzed. The researchers considered whether the conceptual model provided a good fit to biases reported by higher education faculty.

Subjects
One-stage cluster sampling was employed to select 54 universities from a national list of 108 institutions offering doctorates in higher education leadership. Faculty teaching in higher education programs were surveyed exhaustively at these institutions. The estimated size of the faculty who teach planning was 243 higher education professors, yielding a response rate of 67 percent (N=162), or 40 to 85 percent depending on geographic region.

Results
Magnitude scaling of the forced-choice triads (ALSCAL, SPSS-x, edition 2.1) disclosed two dimensions of response. One pattern could be associated with a tendency to emphasize interest in the human resource dimension of planning and the stress of change on organizational values. This orientation also tended to downplay interest in the technical aspects of planning. The opposing pattern stressed a rational and goal-oriented emphasis, emphasizing systematic monitoring, though not necessarily discounting the impact of change on stakeholder interest and organizational values. Check-list responses indicated rival preferences for either qualitative and illuminative data, or for technical approaches and measured feedback. Content analysis of instructional resources revealed that respondents preferred authorities and texts that explicitly analyzed planning as a discipline; or they preferred domain specific information that indexed planning to specialized disciplines (e.g., adult education).

Implications
What emerged appeared to be diverse examples of implicit planning curricula. Faculty might stress brow and specialized domains, or treat planning like a formal domain. Opinion about planning concepts and skills could be humanistic and intuitive, or rational and technical. The management of change, however, requires an integration of the humanistic with the formal, as well as the intuitive and the rational in order to mediate between levels of organizational structure. The researchers suspect that heuristic planning strategies, those that foster the

management of complex problems, were not explicitly articulated in the programs they studied. Further discussion is needed to explore these contradictions.

References

Cope, R. (1981). Strategic Planning, Management, and Decision Making. AAHE - ERIC Education Report Number 9, American Association for Higher Education, Washington, DC.

Cope, R. (1987). Strategic Planning in Higher Education, Opportunity from Strength. ASHE - ERIC Education Report Number 8, American Association for Higher Education, Washington, DC.

Miller, J. (1983). Strategic Planning as Pragmatic Adaptation. ASHE - ERIC Education Report Number 9, American Association for Higher Education, Washington, DC.

Morrison, J., Renfro, W., & Boucher, W. (1984). Futures Research and the Strategic Planning Process: Implications for Higher Education. ASHE - ERIC Education Report Number 9, American Association for Higher Education, Washington, DC.

THE FUNDES-CASE: A MODEL FOR THE DEVELOPMENT OF INSTRUCTIONAL DESIGN

C. Terlouw, Educational Centre, University of Twente, Enschede, The Netherlands

The theme of the symposium will be introduced by means of a model for the development of instructional design (see above). This model is illustrated with a case from an instructional design project in higher education. The different symposium papers will be placed to give insight in the aspects which will be covered by the symposium.
Instructional design in higher education is not a stand alone activity which takes only the direct instructional needs into account. Instructional design has to meet the following factors which have a characteristic realization in higher education:

Developments in the surrounding society
Higher Education can rejoice in an increasing interest. This concerns not only secondary school leavers but also adults who belong to such groups as elder or even retired people, cultural minorities, women, etc.. Therefore, different motives are under discussion when people ask for admittance to the university: professional training, knowledge about sciences, to pass one's free time, reflection about the own experiences in professional practice, emancipation, restarting the career, refreshing one's knowledge, supplying one's knowledge, etc..

Characteristics of (young) adults
Students are young people growing up. Students have their own characteristics, typical for their stage of development, which must be taken into account. It concerns such problems as the fuzzy structure of their knowledge, the lacking motivation and the change to selfregulation.
In contrast to young adults, adults have a structure of knowledge which is purified by their experience. This can be a problem too iif the experience is not based on much different experience situations. Like the young adults motivation is also crucial for them and usually they are very motivated to learn, especially when it has a relationship with their working experience. The independence of adults is a starting point for adult education.

Characteristics of the knowledge to be learned
Problem finding and problem solving are increasingly the central instructional goals in Higher Education. This means that several kinds of knowledge must be acquired to find and solve the problems in some domain.

1217

Educational technology for problem finding and solving momentary seems to be at the cross roads on which new mindsets and new directions are discussed, competing paradigms are struggling to get the upper hand and current beliefs about educational technology are challenged.
This debate also refer to another kind of assumption: the object of learning, the nature of the knowledge itself.

Characteristics of the situation in which instructional design is applied
Instructional design is applied in some faculty in with some teachers are involved. Different forms of cooperation are possible dependent on what is seen as the goal of de instructional design project. Two extremes are distinguished: on the one side a project in which the scientific development of instructional design theory is the only goal of the endeavour; on the other side a project in which the practical development of instructional artefacts is the only goal. Avoiding the extremes one can try to take a position in which the ID-scientist and the teacher(s) in Higher Education cooperate in a team with a double-task: development of both ID-theory and usable instructional artefacts for practice.
If we accept that instructional design must meet (at least) the above-mentioned factors, research concerning the development of instructional design must study these factors in relationship to instructional design.
A model for this approach of instructional design study is depicted in figure 1

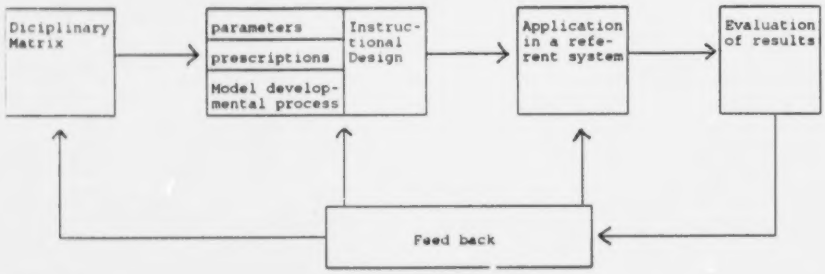


Figure 1
Model for the development of instructional design

The model is primarily based on work of Elen (1991) and a related model of Terlouw (1987). We distinguish several structural components in the design of education in Educational Technology: a disciplinary matrix in which an instructional design model is grounded, an application of the instructional design model to some referent system, the results of the application and its feed back to the former components (Terlouw, 1987; Elen, 1991). A disciplinary matrix concerns the entire constellation of beliefs, values, techniques, assumptions, generalizations and examples which are shared by the members of some scientific community. In this context the beliefs, values, etc. are about how we experience the world, human understanding, the origin of meaning, the nature of knowledge and in particular the nature of human learning and the possibilities for influence by

1218

BEST COPY AVAILABLE

instruction. As said, an instructional design model (from now ID-model) is grounded in some disciplinary matrix. The knowledge base used in an ID-model not only consists of theories on various aspects of instructional and learning processes (Elen, 1991), but also consists of the values, assumptions, beliefs, techniques, etc. connected with these theories. An ID-model contains three aspects (Elen, 1991): first, the design parameters which reflect the theoretical orientation of the model and indicate what variables are considered for instructional design; second, design prescriptions which enable instructional designers to implement the design parameters by operationalizing them in rules and instruments; three, the developmental process which indicate in what sequence the design parameters or how and when the design prescriptions are considered. The referent system specifies the type of instructional-learning situation in which and for which the ID-model is applied. When the ID-model is applied in a instructional-learning situation the resulting design is seen as a practice and theory based hypothesis proposed to an instructional learning problem. This hypothesis must be tested in real practice (artefact as a hypothesis). The results of the test are evaluated and feed back to one or more of the structural components mentioned before.

Terlouw (1987) used this model evaluating the FUNDES-approach. FUNDES is a mnemonic for FUNctional DESign, a procedure for the development of problem solving courses in higher education. The disciplinary matrix concerns the action-theoretical approach based on Vygotsky's ideas and assumptions about psychology and methodology. Therefore, in instructional design, important design parameters are "tool", "orientation", zone of proximal development, etc.. The prescriptions are connected with these parameters: e.g. concrete prescriptions are given to develop a tool to be learned such as a systematic problem approach. All these prescriptions are gathered into the FUNDES-procedure, a model of the developmental process. This ID-model is applicated in the faculty of Public Administration in which two courses are developed, including formative and summative evaluation. The evaluation results are feed back to the courses in the faculty and to the ID-model, especially the prescriptions and the model of the developmental process.

This is one way for a scientific development of instructional design. Other ways are possible too. Therefore, our *general question* is: Which ways can be followed for a scientific development of instructional design?

The papers in the symposium will discuss several methods applied to a concrete subject in higher education or further education for adults:

- an empirical method in which artefacts are constructed and validated based on prescriptions generated from learning and/or instruction theories (Elen, Schott and Podolskij);
- a philosophical method in which the assumptions in the disciplinary matrix about the content to be used in instructional design are critically investigated (Korthals; Van der Meer);
- a reconstruction method in which the instructional design rules or principles actually used in a case in some referent system are analysed and reconstructed (Pouw; Gastkemper).

1219

References

Elen, J. (1991). Psychological bases of instructional design models. Paper Symposium "Instructional design": a challenge to cognitive psychology. 4th EARLI conference August 24-28 Turku, Finland.

Terlouw, C. (1987). De FUNDES-procedure in onderwijsontwikkeling. Evaluatie van een procedure van onderwijsontwikkeling voor het leren probleemoplossen. Doctoral dissertation. Universiteit Twente Enschede.

FROM LEARNING CONCEPTIONS TO INSTRUCTIONAL DESIGN: GENERATION AND VALIDATIONS OF PRESCRIPTIONS

J. Elen, Dienst Universitair Onderwijs, University of Leuven, Leuven, Belgium

Learning conceptions give rise to theoretical frameworks and empirical research. These findings indicate what to consider in instructional practice, they do not, however, provide specifications on the precise consequences for instructional practice. In this study it is assumed that learning conceptions and more generally descriptive theoretical and empirical findings cannot directly be translated into prescriptions. Such findings however enable theoretically sound evaluations of existing instruction. Evaluation of (specific aspects of) instruction implies the construction of valid instruments. Application of the instrument leads to the identification of hypothetical problems for specific categories of students. Research on instruction (instructional theory) provides alternatives to problem solution. Research on these alternatives enables the formulation of hypothetical predictions on the effectiveness of each alternative to solve the identified problem and to optimize instruction for specific categories of students. It is assumed that validation of the hypothetical predictions leads to theoretically sound prescriptions. In this study, this line of thought is emplified for one particular problem, i.e. identification of main points in printed course materials at the university level.

Recent findings with regard to learning, cognition and instruction seriously challenge traditional instructional design practice in its theoretical underpinnings. The parameters considered in instructional design are replaced or interpreted differently. Valuable efforts in research on learning, cognition and instruction to investigate relevant and practice oriented questions, need completion with efforts to make the outcomes applicable in the design of instruction. In this study instrumentation of research outcomes is regarded as one step further to close the gap between description and prescription. A procedure is proposed to generate and validate prescriptions. Descriptive findings cannot be directly converted into prescriptions but they enable theoretically sound evaluation of existing instruction. Evaluation of (specific aspects of) instruction implies the construction of valid instruments. Application of t he instrument leads to the identification of hypothetical problems for specific categories of students. Research on instruction prov ides offers alternatives to problem solution. Research on these alternatives enables the formulation of hypothetical predictions on the effectiveness of each alternative to

solve the identified problem and optimize instruction of specific categories of students.
In our study, this procedure is exemplified for one particular problem, i.e. identification of main points in printed course materials at the university level. In order to detect hypothetical problems of students with indicating the importance of Clauses the following method was applied. The authors of a coursebook on medieval history indicated the importance of each clause of the first chapter on a five-point scale. The same chapter was analyzed by applying Meyer's procedure. It was hypothesized that a discrepancy between
the importance level of a particular clause in the resulting text structure and the rating of the authors, would cause problems for students to indicate the importance of that clause. Three types of clause can be distinguished using the discrepancy scores:

- Clauses high in the text structure but with a low importance rating;
- Clauses low in the text structure but with high importance rating;
- Clauses high (or low) in the text structure and with a high (resp. low) importance rating.

It was hypothesized that students would experience difficulties in rating the importance of (1) and (2) Clauses but not with (3) Clauses. Printing main points bold, adding a summary with main points and text restructuring were investigated as alternatives for supporting students. Prior knowledge, cognitive strategies and metacognitive strategies were regarded to be the main determiners of learning from printed materials.

Method

Subjects and materials

The study was carried out with 84 first year history students at the university level. Data were gathered on students' prior knowledge (test constructed in collaboration with the authors), cognitive and metacognitive strategies (the Inventory Learning Styles developed at the University of Tilburg was used).
Starting from the original version of the chapter, three additional versions were constructed. In each of these versions a different
instructional treatment was used to support students in identifying the importance of information-elements. In the bold-version, important points (highest rating by the authors) were printed in bold, in the summary version a summary with the main points was added
to the text, and in the restructured version the text structure itself was changed in order to realize a better compatibility between text structure and importance ratings by the authors.
Immediately after studying (one version of) the chapter and 23 days later, students rated the importance of 30 Clauses from the text
and answered a post-test (two post-tests were constructed in collaboration with the authors).

Procedure

Three sessions were organised. The first and third session took place during regular lecture times. In the first session, the study was introduced and prior knowledge test and I.L.S. were administered. The second session (one week after

the first one) took place during students' free time. In this session students were asked to study one of the versions of the chapter (distributed at random). After studying a version of the chapter, students rated the importance of 30 clauses from the chapter. Next, the post-test was administered. In the third session students were similarly asked to rate the importance of a 30 ,different from the one in the second session) clauses and answer the questions on a delayed post-test.
In order to gather qualitative data on the strategies used to determine the importance of text elements, the prior knowledge test an d I.L.S. were also administered to an additional group of students (N=10). Every student of this group studied a version of the chap ter. Immediately after having rated the importance of clauses and answered post-test questions, these students were interviewed.

Results

With regard to the indication of the importance of information-elements, no differences were found between the four versions of the chapter (one-way ANOVA) immediately after studying nor 23 days later. Moreover, no interaction-effects were found between student characteristics and text versions as far as the importance-rating of information-elements is concerned. In order to increase the power of the analysis, ANCOVA-procedures were applied. No effects of text version were found.
With regard to the post-tests similar results were found. Only non-significant differences were detected between the results on the post-tests (both immediate and delayed) of students studying different text versions. No effects of text version are found even after application of ANCOVA-procedures. However, significant correlations (1%-level) were found between students' scores on the prior knowledge test and the post-tests. Students with high prior knowledge outperformed students with lower prior knowledge on (immediate and delayed post-tests). Moreover, (significant positive and negative) correlations were found between specific cognitive and metacognitive strategies and results on (either reproductive or productive questions in the) post-tests. For the immediate post-test, these correlations indicate that students scoring high on the 'concretizing' and 'self-regulation' have lower results on reproductive questions, whereas students with high scores on the 'external regulation'-scale have lower results on both reproductive and productive questions. For the delayed post-test, positive correlations are found between scores on the 'analytical' and 'concretizing' scale of the I.L.S. and resp. results on reproductive and productive questions. Furthermore, scores on the 'external regulation'-scale correlate negatively with results on the delayed post-test.
In general, the retrospective interviews indicated that students have strong personal conceptions of what is important in a history text at the university level. One text-based strategy (amount of information on a topic) was reported. Explicitly cueing important information-elements, however, did not challenge personal conceptions.

Discussion

The results of the empirical study may appear to be rather disappointing. None of the theoretical predictions were validated, i.e. none of the instructional interventions supported students to better indicate the importance of clauses.

Various reasons can be mentioned to explain these results: number of students, difficulty level of the text, student capability, inadequacies of the instruments used. However, none of this explanations seems to be valid if confronted with the real data. Prior knowledge tests and post-tests have high reliability scores, the I.L.S. has proven to be a valid instrument in several settings and indicates a wide variety in student's use of cognitive and metacognitive strategies. Therefore, it seems that another explanation has to be looked for which is very valuable for predictive research on instructional interventions. The instructional interventions tackled only one aspect of the learning environment. Data from the retrospective interviews however indicate that students use the global environment to determine the importance of clauses. There fore, rather than trying to formulate prescriptions directed to solving specific problems with regard to learning in instructional settings, it seems to be more valuable to alter the global environment itself. It also indicates the need for comprehensive theories of learning and instruction that can constitute the basis for the design of powerful learning environments.

GENERIC TASKS TO FOSTER LEARNING, TRANSFER AND TESTING IN HIGHER EDUCATION

F. Schott, Deutsches Institut für Fernstudien, University of Tübingen, Tübingen, Germany

In this paper the nature and the application of a tool of instructional design - generic tasks - will be discussed. "Generic tasks" consist of the structural elements of a task to be learned by which a class of concrete tasks is defined. The application of this tool in a concrete case will be reported too.

PSYCHOLOGICAL THEORY AS BASE OF INSTRUCTIONAL DESIGN AND AS PART OF CURRICULUM IN POSTGRADUATE EDUCATION

A. E. Podolskij, Psychological Faculty, Moscow State University, Russia

There exists one sphere of postgraduate education where theoretical-psychological fundamentals of curriculum and teaching process design appear to be simultaneously an essential material which should be assimilated by students to apply in future practice. A case in point is the postgraduate courses of industrial instructors' psychological preparation and training as far as of educational psychologists', concerning with modern instructional technologies. These two kinds of postgraduate students certainly seem to be very different both from objects' and from subjects' point of view. Nevertheless there exists the mutual key point which determines successfulness or failure of such courses. We mean a contradiction between students' past experience, professional attitudes, on the one hand, and new instructional technology' demands, on the other hand. Principles of

psychological theory of "stage-by-stage formation of mental actions" by Russian psychologists P. Galperin were used in two manners:

- as a base to design a postgraduate course and to overcome the contradiction mentioned;
- as a main theoretical (although practically oriented) content to be learnt by the students.

Practical outcomes of such approach as far as some theoretical problems are discussed. A special attention is paid to the dramatic dynamics of change of instructors' primary technocratic requirements.

MATERIAL AND FORMAL EDUCATION IN THE UNIVERSITY, THE CONSEQUENCES FOR INSTRUCTIONAL DESIGN

M. Korthals, Faculty of Social Sciences/Dept of General Education, University of Nijmegen, Nijmegen, The Netherlands

The traditional, neo-humanistic, idea of education in the University presupposes an encyclopedic conception of knowledge, an Unitarian idea of culture and the concept of an University-trained person as legislator of culture. Nowadays these presuppotion, are questionable. There are plausible arguments for a conception of fallible knowledge, a pluralistic, non sexist concept of culture and a more modest concept of the educated person as interpreter of culture. In the philosophy of higher education these new conceptions are until now not sufficiently enough considered. In general, the reaction to changes in the University consists in reviving some kind of idea of Bildung, which implies material education in the sense of knowledge and valuetransmission. Some propose core curricula with established cultural items, other argue for lists of great books. For example Bloom's, Derrida's of Gadamer's theories, but also the educational proposals of the College de France (Bourdieu) belong to this type of enforced revival. Here it will be argued that all these types of material education are seriously flawed. Habermas and Young have argued for formal education, in the sense of that only participation in discoursed can educate students to postconventional intellectuals. They reject value transmission and emphasize formal learning abilities. But mere formal education is insufficient, because it takes not into account the intellectual, moral and affective development of adolescents. In using the concept of categorical education of the German philosopher Klafki, I will argue for a third position which does not neglect student's learning processes with respect to claims of knowledge, competing claims of cultural heritages and a general level of societal education. The consequences for instructional design will be discussed.

A NEW CONCEPT OF SCIENCE AND UNIVERSITY EDUCATION AND ITS CONSEQUENCES FOR INSTRUCTIONAL DESIGN

L. v.d. Meer, Dienst W & S, University of Amsterdam, Amsterdam, The Netherlands

In the nineteenth century, universities began doing research, but did not loose their traditional functions, such as transferring intellectual heritage, promoting personal development and improving the welfare of society. But, mainly due to the character and influence of the modern natural sciences, university education has progressively turned into a specialised vocational training in doing research, whereby the personal, moral and more generalized aspects of expectation have been minimalized. University education is now almost totally in the service of research institutions, and has lost its own character, which is a source of urgent pedagogical and didactical difficulties in university education. A new concept of university research must be developed, which encompasses the personal development of the student and the societal orientation of education. It is necessary to develop a new theory of university education. The consequences of this approach for instructional design will be discussed .

INSTRUCTIONAL DESIGN FOR COMPUTER ASSISTED LEARNING IN HIGHER EDUCATION

C.L.M. Pouw, Educational Centre, University of Twente, Enschede, The Netherlands

Introduction

Instructional design for computer assisted learning (CAL) concerns at least three levels: the lowest level on which the concrete interactions are constructed (e.g. a simulation, a question), an intermediate level on which the nature and sequence of interactions are planned within one or more lessons and the highest level on which measures have to be taken to integrate computer assisted learning in a course or in the curriculum to be sure of an adequate implementation. The different levels set constraints for each other and an instructional design procedure for CAL must take these interrelated constraints into account. In this paper we will discuss such a procedure based on the results and reflections concerning a CAL-project executed at the University of Twente about a computer assisted practice for quantum mechanics.

Theoretical framework

The theoretical framework is primarily based upon a model of educational development (Reigeluth, 1983) and more specific of a procedure for the development of problem solving education, the FUNDES-procedure (Terlouw, 1987). This functional design procedure is a methodology for instructional development based on the activity theory as specified by Van Parreren (Van Parreren, 1988). The research question to be answered is to what extent the rules

for the design of a whole course, especially from the 'design' phase, which is in this case the relevant phase from the FUNDES-procedure, can be applied as rules for the instructional design of a CAL-program.

Methods

The investigation concerns a single-case study (Yin, 1984) of the design of a computer practice in a physics course at the University of Twente about quantum mechanics. This is an example of embedding a simulation for educational purposes in a learning environment. The quantum mechanics course consisted of a series of nine lectures, a text book and exercises. There was also a simulation program available which enabled the students to solve numerically the schrödinger equation, which is an essential differential equation in basic quantum mechanics, in situations which could be chosen freely. The program has been offered to students in a separate practice session at the end of the course. Although the students highly appreciated the simulation program, there were a number of indications that they did not profit from it maximally, as is confirmed in other investigations (e.g. Njoo, 1991). The decision was made to design a new practice for the course. In the present investigation we apply the FUNDES-procedure as a framework to the development history of this new CAL-program. This procedure has been chosen because there was a lot of experience in applying such a procedure in the development of CAL for learning problem solving methods in physical sciences. We will report to what extend this procedure has been realised or not and whether or not it is applicable to the design of an educational environment for a simulation program.

Results

In this summary we omit the case history in detail, which is documented elsewhere (Pouw and Terlouw, 1991). Here we directly apply the theoretical framework to the historical facts. We concentrate on the second step in the FUNDES-procedure, the phase of the DESIGN. Within this phase 8 substeps or actions are suggested (Terlouw and Pilot, 1989):

- Evaluating the existing course:
  Earlier investigations showed that the students did not reach the level of abstraction that was desired (Pouw, Pilot and Terlouw, 1989). Evaluation of the existing course, with the simulation practice on the end of the course, showed that it did not contribute to the objectives of the course. The first question to be answered was whether and to what extend the computer program could attribute to a better or faster understanding and learning of the subject matter involved.
- Choosing relevant teaching-learning theories:
  For the design of an educational environment we have token a learning theory in which instructional functions are defined as general operations or actions that have to be performed in instruction to realize the objectives. These functions are denoted as orientation (presenting information and making the information usable by means of relations and examples), exercising and practising (applying knowledge and giving feedback during practice), testing the learning result and giving feedback after testing. The best way to realize an instructional function depends very much on the specifics and context of the course.

- Formulating educational objectives:
  The main objective for the application of computer assisted learning was to offer students the possibility to get experienced with relations between abstract concepts by means of simulations. In the quantum mechanics case the frame is formed by the idea that the learning of abstract (physical) concepts may be improved by an explicitation of the relations within a network of concepts, mathematical formulas and the physical reality.
- Determining the domain of contents:
  This phase was executed more or less parallel with the third phase. It contains concepts and relations between concepts, theories, laws and mathematical formulas. Two subject matter experts have been observed and interviewed while they were practising with the simulation model.
- Developing a normative model of activity:
  This action, which is characteristic for problem solving programs, has not been performed here.
- Formulating guide-lines from (5) for student learning:
  The guide-lines were not derived from (5), but from (4) in combination with a system of heuristics about learning abstract physical concepts (Pouw, Pilot and Terlouw, 1989). Let the students be involved interactively in visualising the consequences of a model in different physical situations which have been initiated by themselves. Let the students interprete the results of a simulation in relation to the mathematical model and/or the physical reality by means of explanations and predictions. The role of the subject matter experts was predominant here.
- Deducing instructional functions from (2):
  In order to strengthen the exercise function of the simulation the functions of orientation and orientation afterwards had to be realised in the CAL-program.
- Designing a model of the teaching-learning process:
  The structure of the sequence of the type of interactions in relation to the instructional functions was designed. This action has directly been coupled to a first construction of the CAL-program, which can be denoted as 'constructive prototyping'. This action is missing in the DESIGN phase and it goes on ahead the third main phase CONSTRUCTION of the FUNDES-procedure. There is a more or less cyclic phase of prototyping and redesign, in which the cooperation between the CAL-designer and the subject matter expert and the application of an authoring system are essential.
  After that the next main phases of FUNDES (construction,execution/formative evaluation, summative evaluation and decision) followed, but they are out of the scope of this paper.

Discussion and conclusion

Although we performed a single-case study, this is an exemplary one for designing CAL in higher education, especially as far as it concerns the implementation of a simulation within a course. At least three levels of design have been discerned. The most important level of the design is the level of the course: what are the objectives of a course, how can they be reached and what could be the role of the application of computers in it? By following the DESIGN actions 1 to 4 of the FUNDES-procedure in the development of the CAL program all of these general course aspects have been drawn into the CAL-program.

1227

The second level of design has to be strongly related to the first one. This level has to be based on the learning strategy you want to apply (phase 6 to 8). For example a lot of CAL programs have been designed on the basis of a learning strategy of a systematic problem solving approach. In our program it is based on a set of heuristics about learning abstract concepts by explicitation of relations, in combination with a functional approach.
Thus we are reaching the lowest level of design, the direct interactions between student and computer program. This phase of 'constructive prototyping' is not mentioned in the FUNDES-procedure. The types of interactions which have been applied in the quantum mechanics program are questions which have to be answered in textual form, multiple choice questions and the use of simulations in which students can vary parameters according to his/her own choice. The choice of type of interaction depends on the phase in the learning process.
As a result the program consists of simulations embedded in an educational environment in order to realise the educational functions and with that the objectives of the practising sessions. Thus by following the DESIGN phase from FUNDES with some modifications offers a basis for the instructional design of a simulation program as a substantial part of a course.

References

Njoo, M., &, Jong, T., de (1991). Support for learning with computer simulations. Paper AERA Annual Meeting April 3-7. Chicago.

Pouw, C. L. M., Pilot, A., &, Terlouw, C. (1989). Leren handelen aan abstracte, quantummechanische begrippen. Paper Symposium "Ontwerp en evaluatie van probleemoplosonderwijs". ORD 89 May 24-25. Leiden, the Netherlands.

Pouw, C. L. M., &, Terlouw, C. (1991). Het leren handelen aan abstracte begrippen uit de quantummechanica met een computerondersteund praktikum. In L. F. W. de Klerk (ed.). Onderwijsleerprocessen:cognitie en motivatie. ORD 91 May. Amsterdam.

Reigeluth, C. M. (1983). Instructional design theories and Mmdels. Hillsdale: Lawrence Eribaum.

Terlouw, C. (1987). De FUNDES-procedure in onderwijsontwikkeling thesis. Enschede: University of Twente.

Terlouw, C., & , Pilot, A. (1989). Teaching problem solving methods in physical and social sciences. Paper 2nd EARLI-Conference September 4-8. Madrid.

Parreren, C. F., van (1988). Ontwikkelend onderwijs. Leuven: Acco.

Yin, R. K. (1984). Case Study Research, design and methods. Beverly Hills: Sage Publications.

THE DEVELOPMENT AND IMPLEMENTATION OF A PROBLEM–BASED CURRICULUM

E. de Graaff, Delft University of Technology, Delft; P.A.J. Bouhuijs, State University of Limburg, Maastricht, The Netherlands

Introduction

Problem–based Learning (PBL) is thriving in higher education all over the world. During the past decade many institutions in various fields have adopted this educational method. The success of PBL is underscored by superlatives like: "Problem–based Learning (PBL) is the most important innovation in education for the professions for many years; some argue that it is the most important development since the move of professional training into educational institutions" (Boud & Feletti, 1991). Yet the same authors also state that PBL is controversial. Claims of success have been met with skepticism from traditional schools and those who want to introduce PBL in an existing school are advised to think twice before they start (Schmidt, 1990).

There is indeed good reason for careful consideration before starting to change an existing educational programme. The development and implementation of a new curriculum is a large scale operation involving many persons. The process is even more complicated when it involves a major change in educational method. The introduction of a new curriculum based on the method of Problem–Based Learning in an existing school may indeed present many problems. Next to the "normal" problems of curriculum innovation the conversion into a problem–based curriculum concerns a fundamental change in the way of thinking about teaching and learning.

Originally PBL has been developed in the context of medical education at McMaster University in Canada. The approach is characterized by the use of practice related problems to engage students in study activities (Moust et al, 1989). Traditional disciplinary boundaries are often crossed. Students have to deal with the problem on their own, working together in small groups, instead of listening to lectures. Groupteachers (tutors) do not have to be experts in the field. Along with the growing number of schools professing PBL, practice covered by this label has become more varied. There are many examples of curricula who adopt some but not all of these features (Boud & Feletti, 1991). For instance in Maastricht the absolute banning of lectures is abandoned. A taxonomy of different approaches of PBL is presented by Barrows (1986).

1229

The design of a problem–based curriculum implies a series of choices in relation to the professional domain and the actual situation at hand. The first paper of the symposium gives an outline of the early stages of the implementation of a problem–based curriculum at the Department of Building Sciences of the University of Technology in Delft, introducing a series of papers on aspects like: group–size and instructional techniques, faculty development, assessment of learning results, program–evaluation, the introduction of computer assisted learning, and curriculum management. One paper discusses the somewhat different approach with the introduction of a problem–based curriculum at the Department of Management Sciences at the University of Nijmegen.

References

Barrows, H.S. (1986). A taxonomy of Problem–based learning methods. Medical Education, 20, 481–486.

Boud, D., & Feletti, G. (1991). The Challenge of Problem–based Learning. London: Kogan Page.

Moust, J.C.H., Bouhuijs, P.A.J., & Schmidt, H.G. (1989). Probleemgestuurd leren [Problem–based learning]. Groningen: Wolters–Noordhoff.

Schmidt, H.G. (1990). Onderwijskundige aspekten van probleemgestuurd onderwijs. In W.S. Jochems (red.), Activerend Onderwijs, over onderwijsvormen die het leren bevorderen [Activating Education, on educational formats that encourage learning]. Delft: Delft Universitaire pers.

THE INTRODUCTION OF PROBLEM–BASED LEARNING AT THE DEPARTMENT OF BUILDING SCIENCES

E. de Graaff, Delft University of Technology, Delft; P.A.J. Bouhuijs, State University of Limburg, Maastricht, The Netherlands

Introduction

A few years ago the department of Building Sciences of the Delft University of Technology was confronted with severe criticism on the educational program (Verkeningscommissie, 1989). The department recognized the need for thorough restructuring of the curriculum and appointed the Program Committee Building Sciences (PKB) with the assignment to generate within the year recommendations for a new curriculum with the following characteristics: better numerical efficiency, technological emphasis in the basic years and a more differentiated graduate profile. Faced with the task of designing a new curriculum in such a short time the PKB requested educational advice from the Educational Research and Development department of the University of Limburg.

Besides the criticism on the content of the curriculum the major problems were a lack of integration of knowledge from different domains and the substantial study delays, amounting to only 22% of the students completing their studies within the maximum allotted time of six years (Statistisch jaarboek '88/'89). Integration of disciplines is one of the main features of problem–based learning (Boud & Feletti, 1991), and the problem–based Maastricht medical curriculum has a high

numerical efficiency compared to other Dutch medical schools (Post et al, 1988). Problem–based learning appeared to be a way to tackle these problems and at the same time provided a clear signal to the outside critics that a substantial change was set in motion. The PKB proposed to design a new curriculum, based on the principles of problem–based learning, with more emphasis on technic and with two new graduation majors (PKB, 1989). The proposal of the PKB was accepted by the Faculty Council as a foundation for the development of a new curriculum by the end of 1989. The first students actually entered the new program in September 1990.

Problem Based Learning

Problem–based learning (PBL) has been developed within the scoop of medical education at the University of McMaster in Canada. Actually PBL is not just a set of educational techniques, but rather an educational philosophy, an approach to learning (Neufeld & Barrows, 1974; Engel, 1991). Essential is the shift from teacher centred thinking to student centred thinking, from a discipline oriented curriculum to one that intends to promote learning.

Important features of PBL are (Moust et al, 1989):

- integration of disciplines and skills;
- curriculum structure with thematic blocks;
- preparation of the blocks by teams of teachers;
- learning oriented work in small groups;
- self–directed learning.

The Building Sciences version of PBL

Building Sciences and especially architecture has an educational tradition which cherishes the old values of tutelage and apprenticeship to a practitioner (Maitland, 1991). The Delft school of architecture has gained international fame with the model of project education developed during the sixties. Integration of knowledge and skills was one of the important objectives of the model of project education. A strong argument in favour of PBL was that it offered an opportunity to renew the attempts to achieve this objective of integration.

The proposal for the Building Sciences PBL–curriculum was modeled after the Maastricht example. Students are offered opportunities for learning rather than a one way stream of information. The curriculum is developed according to centrally defined principles. The thematic outline for the first two years of the programme was designed by the PKB. The two year basic curriculum consists of 12 six–week thematic blocks (with themes like: The House, The City, Function and Form and Renovation). The study–programme for each block is prepared carefully by a team of teachers, laid down in a blockbook and appraised in advance by a programme committee.

Further developments

Important themes in the first two years of implementation have been professionalization of the staff, especially the training of tutorial skills, the development of study facilities within the context of problem–based learning, the development of assessment methods to be used in a problem–based curriculum,

the design of a systematic approach to programme–evaluation and the central coordination of the educational programme.
By now the development of the new Building Sciences curriculum is firmly on its way. The blocks of the first two years have been constructed according to plan. Several adaptations have been made along the road and more are to follow. Building Sciences is now heading for a truly own version of PBL with as main characteristic the integration of knowledge and skills in the design education.

References

Engel, C.E. (1991). Not just a method but a way of learning. In D. Boud & G. Feletti (eds.), The Challenge of Problem–based Learning. London: Kogan Page.

Maitland, B. (1991). Problem–based learning for an architecture degree. In D. Boud & G. Feletti (eds.), The Challenge of Problem–based Learning. London: Kogan Page.

Moust, J.C.H., Bouhuijs, P.A.J., & Schmidt, H.G. (1989). Probleemgestuurd leren [Problem–based Learning]. Groningen: Wolters–Noordhoff.

Neufeld, V., & Barrows, H.S. (1974). The Mc Master Philosophy: An approach to medical education. Journal of Medical Education, 49, 1040–1050.

Programma Commissie Bouwkunde, PKB (1989). Herprofilering van de Bouwkundeopleiding aan de Technische Universiteit Delft; [A New Profile for Building Sciences at the TUD; Foundations of a new and broadened educational programme]. Delft: Technische Universiteit Delft, Faculteit der Bouwkunde.

Statistisch jaarboek '88/'89 (1989). Delft: Technische Universiteit Delft, Bureau van de Universiteit.

Verkeningscommissie Bouwkunde (1989). Eindrapportgae [Final Report].

FACULTY DEVELOPMENT AS A TOOL IN THE IMPLEMENTATION OF EDUCATIONAL INNOVATION

Jan van Driel, Faculty of Philosophy and Humanities, Educational Development Unit, Delft University of Technology, Delft, The Netherlands

Introduction

Research in the field of curriculum innovation always shows that innovations introducing changes in the way of teaching can only be successful if teachers actually succeed in performing the desired teaching skills. The implementation of a problem-based curriculum in an existing educational environment places great demands in this respect: teachers are not only asked to change their style of teaching, but also a specific attitude towards teaching is required from them (Schmidt, 1990; Todd, 1991). In short this involves a shift from teacher-centred education to self-directed learning. Taking the case of 'Building Sciences-Delft' as an example, this paper describes a way of promoting attitude development amongst staff, results obtained with the approach chosen and conclusions with respect to the future.

Specific teacher-roles in a problem-based curriculum

When the Dept. of Building Sciences in Delft chose to implement a problem-based curriculum, it was clear from the start that at least two 'new' teacher-roles were introduced (de Graaff & Frijns, 1990). First, because block planning groups were formed in order to develop multidisciplinary, thematic blocks, the role of *block planning group member* was created. Block planning groups include teachers representing the different disciplines present in the block. Working in a team, these teachers had to select subject-matter and adequate instructional materials and methods. Besides they had to write study assignments, cases and tasks. All this had to fit within the framework of problem-based learning (PBL). When they started, most block planning group members had little knowledge about PBL, nor had they any experience in designing PBL-blocks and in making an educational design within a team.

The second new teacher-role concerned the *tutor*, a characteristic feature of problem-based curricula. A tutor is a teacher who supervises groups of 8-12 students, who meet twice a week during 2-hour sessions, led by one of the students (so-called 'tutorial groups'). Each meeting includes a group discussion on a problem, mostly originating from professional practice (a so-called 'case'). This discussion should lead the students to formulating learning objectives, which they should try to attain by means of self-study after the meeting. The next meeting starts with an oral report of the results of the self-study activities. Next, a new case is discussed, etc. (see also Moust *et al*, 1989).

The task of the tutor mainly is to support the students in a process that can be described as *'learning to learn'*. In this respect, the tutor teaches the students to handle problems systematically, to formulate adequate learning objectives and to report orally in a brief and well-structured way. The tutor also stimulates self-study activities by discussing questions like 'how can certain knowledge be found?' and 'how can it be studied efficiently?'. From this description it should be clear that a tutor not only needs specific skills, but also, and more important, a certain *attitude towards teaching*, differing from those connected with 'traditional' teacher-roles and skills.

Method

It was decided by the board of Building Sciences that future block planning group members and tutors should be prepared for their new tasks by means of a training programme. Block planning group members had to participate in two conferences, followed by a tutorial training module of two days. The latter was also presented to future tutors. In fact, all staff with an educational task was obliged to participate in this training module. Finally, block-specific training was offered by the organization of weekly teacher meetings during the six weeks in which a block takes place.

The *conferences* served as an introduction to the principles of PBL and the premises of the new curriculum to be implemented. Also a beginning was made with the design of such a curriculum. The conferences included lectures, mostly held by experts in the PBL-field invited as external consultants, exercises in writing study assignments and discussion sessions.

The main purpose of the *two-days tutorial training module* was to practice tutorial skills. The training module was designed according to principles of PBL, including self-study activities and group-discussions. Arranged in groups of 8-12,

participants simulated tutorial groups, discussing cases acting as though they were a group of students with one of them performing the role of tutor. Videofragments of tutorial groups were used to demonstrate and discuss tutorial skills.

Results

The conferences showed that the principles and premises of a problem-based curriculum were far from self-evident for the staff. Ample discussions promoted understanding however. The most important output of the conferences was that they helped to create an atmosphere in which future block planning group members were willing to work together in a team designing and developing the blocks. They only became able to actually carry out this task after having participated in the tutorial training module. Although this was not the main purpose of the course, it taught block planning group members amongst other things how to design a block (connecting subject matter with educational methods and appliances), how to write cases, etc..

For those who had not participated in the conferences, the tutorial training course served mainly as an introduction to PBL. Especially the simulation exercises led to profound discussions concerning the pros and cons of the new educational system. This partially prevented however tutorial skills to be practised as was the main objective of the course. This training module thus resulted in distinguished learning outcomes for both types of participants, but these were not in the first place in the field of tutorial skills.

Shortly after the start of the new curriculum in September 1990 problems of various nature arose. One type of problems was connected with *imperfections in the blocks:* lack of internal coherence, students were overtasked, inadequate instructional materials, unsuited cases, etc..

Besides performing the role of tutor presented a problem to a lot of staff members. For some of them this had 'merely' to do with a lack of tutorial skills. For others however *a lack of motivation* was most important. One argument frequently used against the tutor role referred to a loss of autonomy: tutors have to function within the framework of an educational system of which the choice was not influenced by them; within a block which they did not design themselves and in a role in which their professional expertise doesn't seem to be recognised. Amongst others these problems were noticed in the weekly teacher meetings. These meetings were visited by only half of the teachers at a maximum. This clearly indicates a lack of involvement.

To meet with these problems, it was decided to organise *workshops after every block* for the staff members concerned. In order to obtain maximal attendance all potential participants were addressed personally. In addition to an evaluative purpose, aimed at adjusting the block, these workshops were also meant to get staff members better acquainted with the new teacher roles. Therefore experiences and problems from educational practice were exchanged and discussed. To emphasize this aspect, the workshops were prepared and led by a teacher trainer. Although participants were enthusiastic about these meetings, the large non-response (ca. 50%) is still to be considered precarious.

Conclusions and discussion

For the near future a plan of promoting staff development is proposed. On the one hand *staff management*, based on the principles of Human Resource Management, should be aimed at deploying teachers as much as possible in blocks and in roles according to their affinity and in ways that stimulate their feeling of responsibility. One option in this respect is to select motivated senior students in order to perform as tutors.

Also, *adjusting the teacher roles* is considered. For example, it would be possible to extend the tutor role with skills relating to professional expertise by combining the tutor role with that of a design studio teacher.

Finally, an inquiry is made of wishes concerning further staff development. Our option is to develop various forms of *permanent in-service training*, like:

- block-specific courses (on both subject-matter and teaching skills)
- workshops and/or follow-up meetings on general themes (e.g. methodology of research and design, group dynamics, etc.)
- training at the micro level of education. For example teachers visiting each others tutorial groups and/or being visited by a teacher trainer, followed by feedback sessions.

The activities considered desirable by the department may be combined within a system of permanent staff development, to be integrated with staff management (van Schoot, 1992).

References

Graaff, E. de & P.H.A.M. Frijns (1990). Teacher Roles in a Problem Based Curriculum. Paper at the 2nd International Symposium on Problem-Based Learning, Yogyakarta, Indonesia, October 7-12, 1990.

Moust, J.H.C., Bouhuijs, P.A.J., & Schmidt, H.G. (1989). Probleemgestuurd Leren. Groningen: Wolters-Noordhoff.

Schmidt, H.G. (1990). Onderwijskundige Aspecten van Probleemgestuurd Onderwijs. In W.M.G. Jochems (Ed.), Verslag Symposium Aktiverend Onderwijs (pp. 16-37). Delft: Delftse Universitaire Pers.

Schoot, T.E. van (1992). Een Systematische Aanpak van Deskundigheidsbevordering. Onderzoek van Onderwijs, 21 (1), 12-14.

Todd, S. (1991). Preparing Tertiary Teachers for Problem-based Learning. In D. Boud & G. Feletti (Eds.), The Challenge of Problem-Based Learning (pp. 130-136). London: Kogan Place.

THE ASSESSMENT OF STUDY RESULTS IN RELATION TO THE EDUCATIONAL METHODS

P. Frijns, Department of Building Sciences; E. de Graaff, Department of Building Sciences, Delft University of Technology, Delft, The Netherlands

Introduction

The main purpose of the Department of Building Sciences at the University of Delft is to educate skilled building engineers, who are capable to contribute to the

development of the architectonical science as well as the development of architectonic profession. However, to what extent are the purpose, the ideals and the intentions truly achieved?
An answer to this question is also to be found in what the educational system requires students to do in order to survive and to prosper. Hence, the way students are assessed predominate the study–approach of the students and may be used to achieve the desired learning goals (Newble & Jaeger, 1983; Frederiksen, 1984).
In this way student assessment serves three goals:

- summative evaluation;
- formative evaluation;
- programme evaluation.

For this reason the assessment should be congruent with the educational principles of the curriculum and the assessment should be a continuous process. Problem based learning requires a specific evaluation system. The three most important demands are:

- the assessment should be in line with the principles of self–directed learning;
- the evaluation should be multidisciplinary;
- the assessment should also incorporate the application and use of knowledge (Van der Vleuten & Wijnen, 1990).

The conventional strategy in which individual teachers were responsible for the examinations was no longer appropriate. Assessment activities had to be centralized. In this paper the development of the new evaluation system of the Faculty of Building Sciences is discussed.

The evaluation system

In concordance with the programme it was decided that three competencies should be evaluated separately: knowledge, technical skills and designing. Although For each competency specific procedures are developed, most attention has been given to the assessment of knowledge.

Assessment of knowledge

Although some authors call the measurement of knowledge irrelevant and artificial (Pickering, 1979) it is obvious that knowledge plays a crucial role (e.g. Glaser, 1984; Norman et al., 1985) .
At the Department of Building Sciences a so called block test is used to assess knowledge. Each block test consists of 65 to 125 true/false items. Sometimes the block test is completed with 5 open–ended questions. The choice for the true/false format was motivated by the fact that knowledge can most efficiently be measured using this type of questions.
The content of the tests is based on a blueprint, specifying the number of questions per discipline and per theme. The blueprint is drawn up by the blockgroup and is in concordance with the goals of the specific block.
The kems for the block test are constructed by the members of the block group. Each item is screened by the Central Test Review Committee; a multidisciplinary team of seven persons which checks the content, the wording, the designing and the relevance of the items. If the committee considers alternation necessary, one of

the members contacts the author of the item. At last, the items are administered in the test.
At the end of each block period all students take the test. Afterwards, the students are encouraged to supply comment on items if the test within five days. These comments together with the psychometric information are discussed with the blockgroup. In case one or more items are disputable, these items are eliminated from the test. The definite results are determined on the remaining set of items.

Technical skills
Within the curriculum of Building Sciences a broad variety of technical and practically skills are trainec. These basic skills are spread over the 12 blocks of the first two years. The goal of these training is to prepare students for the second part of the study, in which they have to specialize.
The importance of these skills is also reflected in the assessment programme. However, no specific instruments for assessing technical and practically skills have been developed yet. At the moment each block group is responsible for the design of the exercise and the assessment of the students. The assessment instruments used vary from making an exercise to writing a report.

Designing
During each blockperiod the students have to complete one or more design assignments. The written instruction is made by the block group as well as the criteria for evaluating the products. At the end of each blockperiod the designs are evaluated, whereafter a grade is given by the teacher. Since September 1991 the department is striving for minimal two ratings for each product to minimalize the subjectivity of the ratings. However, this is not always possible since the grading of these products is very time consuming.
In march 1992 a study is conducted in which the validity and reliability of the ratings is investigated. The main purpose of this study is to get insight into the rating of designs and to come to empirically justified recommendations for constructing check-lists.

Discussion and conclusion
From the beginning special attention has been given to the assessment of knowledge. The emphasis on instruments for assessing knowledge was due to the fact that in the past no integrative tests had been used whereas most teachers had experience with rating practical skills and designs.
Last two years have shown that the examination system has a few side–effects. First of all, problem based learning emphasize the role of self–management. As a consequence, there is a wide variety of learning paths. A block test at the end of each block period frustrates these activities and elicits a memorising approach. Hence, in this way the student–centred approach is reversed to a teacher–centred approach. At the moment a first experiment with a progress test (Van der Vleuten & Wijnen, 1990) is being conducted. At the same time the quality of the true/false items is being studied and new formats are being tested.
A second side–effect concerns the subjectivity of the ratings. Students are conscious of the fact that the rating can be influenced by several factors. One important factor is the neatness of the presentation. A number of students pay too

much attention to the presentation of their design in order to get a higher grade. As a consequence, they have less time left for studying the literature and they fail the block test.

References

Fredericksen, N. (1984). The real test bias: influences of testing on teaching and learning. American Psychologist, 39, 193–202.

Newble, D.I., & Jaeger, K. (1983). The effect of assessment and examinations on the learning of medical students. Medical Education, 17, 165–171.

Van der Vleuten, C., & Wijnen, W. (Eds.) (1990). Problem-based learning: Perspectives from the Maastricht experience. Amsterdam: Thesis.

SYSTEMATIC EVALUATION OF THE EDUCATIONAL PROGRAM AS PART OF THE INNOVATION STRATEGY

M.P.J.G. Claessens, Educational Development Unit; W.M.G. Jochems, Delft University of Technology, Delft, The Netherlands

Introduction

The purpose of program evaluation is to obtain and to report evaluation results on which decisions about installation, continuation, modification or termination of a program can be based. Literature on curriculum innovation program puts much emphasis on the integration of an evaluation study as part of the implementation strategy (Romiszowski, 1981). As part of the implementation of a curriculum based on the principles of problem-based learning at the Faculty of Building Sciences, the planning of an integral system of program evaluation was started during the first year of the implementation of the new program. In this paper, the evaluation model will be presented, along with a brief description of the Delphi method survey, which is an important part of the first phase of the project.

The evaluation model

Doing evaluation research is, like other sciences, to choose. A few of the things the evaluator has to consider are: which evaluation model is the starting-point for the evaluation study, with what kinds of instruments will the evaluation data be gathered, and so on. In this case, starting-point were the so called "decision oriented evaluation models" (see for an overview of models of evaluation: Popham, 1988), in particular the viewpoints of Lee J. Cronbach (1980, 1982). Program evaluation is in his view inherently linked with the political context of a program. So, in designing an evaluation study not only scientific considerations have to be taken in account, but also the practical and political context are relevant. In his writings he also argues that the evaluator has to serve all the relevant participants in the policy making context. Different groups of participants have different interests, and will have answers to specific questions. To overcome the danger of a becoming a member of one group, the evaluator has to be a "multi-partisan". So, it is likely that an evaluation study becomes a mixture of studies,

rather than a single study. Based on these elements a descriptive model was developed.

In this model, three key aspects can be distinguished
First, the evaluator has to have a firm understanding of the needs for information. To put this in another way, (s)he has to get an answer to the following question: "Who needs what kind of information?" Several methods can be used to inventarize the needs: by interviewing key-informants, through personal communication with participants, or by a more thorough needs assessment survey. Next, the evaluator has to inventarize the sources of information. At this stage, the evaluator looks for all the relevant kinds of evidence evaluation research can obtain. More specific, for every need for information, all possible sources of information will be inventarized. Along with, the evaluator thinks about the ways information can be obtained. In other words, the methods and the instruments of the data gathering will be generated. More basically, at this stage the evaluator decides to undertake a scientific or a naturalistic evaluation study (or both). Finally, after the operational stage of the evaluation study (the data gathering), evaluation data are processed and the results will be reported to the various audiences in a for every audience fashionable way.
In the model two features can be distinguished. First, the evaluator is an intermediate between all the relevant needs and the sources for information. The evaluator operates as a linking agent, because (s)he has an overview of the relevant aspects of the evaluation study. Second, in every evaluation study choices have to be made. Too often, no account has been made for the kind of research questions, the methods and the instruments been used. By listing all the possible needs for and sources of information, the evaluator will be more aware of this.

The Delphi method
According to the model, an evaluation study has to start with an understanding of the needs for information. The method here used to assess these needs was the Delphi method (Delbecq, Van de Ven & Gustafson, 1975; Linstone & Turoff, 1975). The Delphi method was first used in the nineteenfifties as an instrument to predict future developments. Nowadays it is frequent used for many purposes, including as a needs assessment instrument (see Prevoo & Voeten, 1989).
The Delphi method consists of a series of questionnaires. Experts are selected and respond to the subsequent questionnaires. This way a group sort of paper-and--pencil is simulated. By eliminating face-to-face interaction, problems of dominating individuals on group decisions are avoided. Stimulated is also a process of more rational argumentation.
In general, the procedure is as follows. Participants are asked to respond to an unstructured, broad question. In the next questionnaire, all the responses are transformed into structured items. Participants respond to every item on a Likert scale and by a written argument. In the third round, every participant receives feedback, such as the mean response of all respondents, and a summary of the written responses (including reasons why some participants do not view the mean response). After reviewing the written responses, participants again respond. This

1239

process stops when consensus is obtained, or when enough information is exchanged.

Subjects

Delphi participants are not a random sample of the population, but will be selected. In this case, through careful selection of participants, all the relevant viewpoints on quality of the new program were included (from proponents to the new curriculum to opponents to the new curriculum). To avoid too subjective selection, the following procedure was carried out. First, 50 members of the staff were selected. Then, with the random numbers procedure, 35 respondents were asked to participate.

Procedure

The first questionnaire was examined in a pilot study (5 participants). In February 1992, the first questionnaire was sent to the participants. Participants were asked to respond to the next question: "Which aspects do you take into account when you judge the quality of the Building Sciences curriculum?"
At the paper presentation, the first results are presented.

References

Cronbach, L.J. (1982). Designing evaluations of educational and social programs. San Francisco, California: Jossey-Bass.

Cronbach, L.J. and Associates. (1980). Toward reform of program evaluation. Aims, methods and institutional arrangements. San Francisco, California: Jossey-Bass.

Delbecq, A.L., Van de Ven, A.H., & Gustafson, D.H. (1975). Group techniques for program planning. A guide to nominal group and Delphi processes. Glenview, Illinois: Scott, Foresman.

Linstone, H.A., & Turoff, M. (eds.). (1975). The Delphi method: Techniques and applications. Reading, Massachusetts: Addison-Wesley.

Popham, W.J. (1988). Educational evaluation (2nd ed.). Englewood Cliffs, New Jersey: Prentice Hall.

Prevoo, E.G.H.M., & Voeten, M. (1989). De Delphi-methode voor het bepalen van opleidingsbehoeften [The Delphi method as a means for assessing training needs]. Gids voor de opleidingspraktijk, 2, 1-19.

Romiszowski, A.J. (1981). Designing instructional systems: decision making in courseplanning and curriculum design. London/New York: Kogan Page/Nichols.

Worthen, B.R. (1990). Program evaluation. In H.J. Walberg & G.A. Haertel (eds.), The international encyclopedia of educational evaluation (pp. 42-47). Oxford: Pergamon Press.

THE ROLE OF COMPUTER AIDED INSTRUCTION IN A CURRICULUM BASED ON THE PRINCIPLES OF PROBLEM BASED LEARNING

P. Frijns, & P. Barendse, Faculty of Building Sciences, Delft University of Technology, Delft, The Netherlands

Introduction

In a traditional educational programme educators are primarily responsible for decisions about what, how, where and when learning will occur. Students are perceived as relatively passive and subservient recipients in a process controlled by the staff. Hence, in a teacher centred educational programme students are not directly involved in decisions about the content and the process of learning.
A curriculum based on the principles of problem–based learning (PBL) is an example of a educational programme in which the traditional power relations are changed and in which both students and teachers share control. Through building a co–operative learning climate and through formulating and assessing their learning needs students are being helped by taking control of the content of their learning. Self–management is the most important skill for students to take control of the learning process (Moust, Bouhuijs, & Schmidt, 1989; Hammond & Collins, 1991).
In a problem–based curriculum the staff should involve the students to design their own learning plans and they have to encourage them to manage their own learning as well. So, the school has to provide an atmosphere in which the students are stimulated to consulate with others (e.g. peers and teachers) and in which students learn to become more autonomous and confident with challenges. A basic collection of primary educational appliances could also support the self–directed learning activities (Moust, et al., 1989).
The Faculty of Building Sciences in Delft has developed a curriculum on the principles of PBL, about two years ago. From the beginning attention has been given to the designing of a Study Centre. In this place the students have the disposal of a basic collection of educational appliances (e.g. books, videos and courseware). The primary goal of the Study Centre is to support the self–tuition activities of the students.
Although a number of educational appliances are available, the development is still going on. Especially, the development of courseware is being emphasized. This paper concerns the experiences with the construction and the implementation of courseware for the benefit of the Study Centre.

The construction of courseware

Although computer aided instructions (CAI) are relatively new educational tools, they have become an important supplement to traditional educational appliances. In the literature a wide variety of instruments and definitions of CAI are available. In the most including definition every instruction in which computers are applied is viewed as CAI. Since in this paper attention is given to courseware in behalf of the Study Centre a more strictly definition has to be used.
CAI is a type of teacher–independent education in which teaching or testing material is stored in a computer. The learning process takes place by means of

direct interaction between student and computer. The computer programme specifically designed for CAI is called courseware (Van Hees, 1983).
In this definition the role of courseware in supporting self–tuition is accentuated and the importance of direct and adequate feedback is stressed as well. According to the definition the following forms of courseware can be distinguished (Allessi, & Trollip, 1991):

- tutorials;
- drill and practice;
- simulations;
- formative computer–testing.

When the Study Centre was put into use only a small number of simulations and tutorials were available, whereas drill and practice as well as formative computer–testing were lacking. Hence, the faculty was in want of adequate courseware. A number of appropriate computer applications were obtainable at other faculties and universities. Nevertheless, a great number of courseware had to be constructed. Of course, this process of constructing new courseware is still going on.
At the Faculty of Building Sciences the construction and the providing for courseware is the task of the Computer Supporting Group (CSG). When a teacher or blockgroup responsible for the content of a certain educational period is interested in the use of computer aided instruction, the CSG is consulted. The purpose of this consultation is to determine whether CAI is the most appropriate way to achieve the learning goals. Whether CAI is recommended depends on several factors which are included in some kind of a check-list. In general, the CSG recommends CAI in case that both the specific opportunities of the computer (e.g. interactions) will be used and CAI will most probably result in better achievements compared to other educational appliances.
After a positive advice has been given, the most adequate form of courseware has to be choosen. Starting from the perspective that the courseware is mend for the Study Centre, one of the forms described above will be selected. Of course, which form of courseware is most adequate depends on the kind of learning goals that is persuade.
In case no appropriate application is obtainable, the desired courseware has to be constructed. The CSG makes use of the usual strategy for constructing courseware, which consists of five chronological steps. The purpose of the first step is to get insight into the content and the structure of the underlying model. On the basis of this information a global scenario is written. Establishing the final scenario as well as the final structure of the model are the goals of the second step. The third step concerns the development of the the detailed technical design and the first release of the courseware. The last two steps include the testing and implementation of the courseware. The results of the pilot study are used to adapt the courseware to the wishes of the users. The second release is tried out in a more formal setting, whereafter a detailled evaluation is conducted. Again, imperfections in the application are repaired. Finally, the courseware is installed in the Study Centre.

Experiences with the construction and implementation of courseware

Before the reorganization of the Faculty of Building Sciences, only few teachers made use of computer applications within the course. However, the last two years the interest for computers has been growing. Nearly all blockgroups of the first and the second year of the curriculum incorporated computer applications in their course. A number of these applications are available in the Study Centre, whereas the greater part is directly used within practical works. Why are most of the applications not available in the Study Centre?

The answer to this question is simple: "lack of time". Theoretically the strategy for constructing courseware is perfect. However, in practice it is not always possible to go through all steps before the application is used. Usually, the teacher consultates the CSG only a few weeks before the starting of the blockperiod. In this way, it is impossible to construct student proof courseware. Therefore, the CSG has developed a more practical strategy: the first release of the application is nor furnished with an instruction neither with feedback. So, the teacher can apply the courseware in a practical work, if he verbally gives the instruction and the feedback to the students. Next, the experiences are used to construct the instruction and the feedback part of the courseware. This release is tried out in a pilot study, whereafter the courseware is implemented in the Study Centre.

Since the construction of appropriate courseware for the Study Centre is very timeconsuming, this strategy for constructing courseware seems to be more effective and more efficient than the theoretical approach. In this way the applications can be tried out sooner in the construction process through which failings can early be discovered.

Results of the programme evaluation showed that the students are pleased to handle the computer applications in the Study Centre as well as in the practical work. So, the results confirm the usefulness of CAI in a curriculum based on the principles of PBL.

References

Allessi, S.M., & Trollip, S.R. (1991). Computer based instruction: Methods and development. Englewood Cliffs: Prentice Hall.

Frijns, P. De Grave, W., & Wolfhagen, I. (1991). Glossary: a list of frequently used concepts in the field of problem–based learning. Maastricht: Department of Educational Development and Research.

Hammond, M., & Collins, R. (1991). Self–directed learning: critical practice. London: Kogan Page.

Moust, J.H.C., Bouhuijs, P.A.J., & Schmidt, H.G. (1989). Probleemgestuurd leren. Groningen: Wolters–Noordhoff.

Van Hees, E.J.W.M. (1983). Computertoepassingen in het onderwijs: een overzicht. Handboek voor de onderwijspraktijk. Deventer: Van Loghum Slaterus.

MANAGING THE INTRODUCTION OF A PROBLEM-BASED CURRICULUM IN AN ESTABLISHED FACULTY: A CASE STUDY

Wim H. Gijselaers, Department of Educational Research, University of Limburg, Maastricht, The Netherlands; Herman van den Bosch, Faculty of Policy and Administrative Sciences, University of Nijmegen, Nijmegen, The Netherlands

The concept of problem-based learning (PBL) constitutes a radical reform in university education. This concept was initially developed within the context of medical education. Over the past ten years problem-based learning has especially been put in practice in disciplines which have a traditional orientation towards professions, such as law, management and engineering (Goodlad, 1984). During this period problem-based learning has gained the reputation of a far-reaching innovative approach to education (Schmidt and De Volder, 1984).
The concept of problem-based learning is regarded as a valuable attempt to improve the quality of higher education. It is an innovative instructional method which presents various theories as an instrument to understand and explain problems. Since 1988, the Faculty of Policy and Administration Sciences (FPAS) of the University of Nijmegen (the Netherlands) has pioneered a curriculum based on this concept. The principal idea is, that learning should be organized around problems which are related to the profession, rather than around subjects which are centred around academic disciplines. Considerable efforts have been made by this faculty to design a curriculum and employ innovative teaching methods which intend to achieve multi- or interdisciplinary education, to encourage self-directed learning, and to provide students with an adequate background to analyze problems encountered in the professions of policy and administrative sciences.
The pressing concern for improving the quality of higher education leads to the conclusion that there is a definite need for studies on improving university teaching and innovation processes in higher education.
This paper seeks to describe the process through which faculty became aware of the need to develop and implement new ideas in teaching (in particular problem-based learning), and to institutionalize these ideas in an organizational context.

Analytical frameworks for innovation and change in Higher Education
Classical innovation theory heavily relies on the concept of stages or phases in innovations: adoption, implementation, dissemination (Hill & Friedman, 1979; Kozma, 1985). However, Kozma (1985) showed that 1) a considerable overlap and ambiguity exists between these stages, and 2) that innovation is evolutionary (new instructional practices are built on past practices). According to Kozma these stages are not easily demarcable. A clear point of adoption or implementation is rarely discernible. This seems indicative for the most characteristic aspect of instructional innovation: new educational practices are based on past practices. Whenever instructors employ new teaching methods, these methods are embedded in earlier teaching experiences.

An obvious question is which conditions have a major influence on innovation processes. Literature on change and innovation suggests that the most pervasive variable is the unique organizational structure of higher education (Bess, 1984; Kozma, 1985). Kozma (1985) points out that academic organizations are characterized by their perpetual inability to strike a wholly satisfactorily balance between the requirements for individual autonomy and academic freedom on the one hand, and the necessity for organizational efficiency, accountability and control on the other hand.

The organizational lousiness, or lack of instructional accountability, accounts for the personal character of many innovations in higher education. Dependence on a variety of personal preferences of instructors leads to unclear and inconsistent choices in decision processes with respect to educational change. Innovations which outsize the level of individual instructors, for example rearrangement of course contents at a program level, are particularly vulnerable for failure and resistance to change. At a program level it is difficult to operate on a basis of a variety of individual, and therefore inconsistent, preferences which can be described better as a loose collection of ideas than a coherent structure (Cohen, March & Olsen, 1972).

A second condition is external pressure (Lindquist, 1978). Declining student enrollment, external peer reviews, changes in higher education policy and finance, changes in public expectations, are generally seen as relevant external forces. Innovations are apt to occur whenever these external forces grow stronger. They are more important than internal forces to invoke educational change.

A third condition is funding or the availability of additional resources. Innovations require time or other resources. Lack of resources, as Kozma (1985) points out, is one of the most frequently given reasons for not adopting an innovation. Instructional change is only possible to the extent that change doesn't cost anything. When costs are incurred, faculty members have to spent time on obtaining additional resources.

A fourth, and final, condition is that successful academic career largely depend on research efforts. Individual preferences therefore normally lie in the field of research. Research activities are perceived as the most rewarding activities in an academic career. However, implementing innovations implies that faculty have to undertake activities which will reduce time for research. Consequently a feeling of being hindered by educational innovations in the pursuit of academic career will emerge which ultimately leads to rejection of new ideas or minimalism to give shape to these ideas (Oldham & Kulik, 1984).

The analytical framework as described above leads to several practical implications. First, powerful management is needed to create coherent structures in a faculty. Coherent with respect to the intended educational goals, methods and evaluation and assessment. Second, external pressure is needed to legitimize the actions undertaken by the management. External pressure provides management a basis to change the balance between individual autonomy and faculty control in the favour of faculty control. Third, different reward systems, which will also take into account efforts in the field of education may have a positive influence on adopting innovations. Finally, management should try use external pressure for adopting new ideas and to create a faculty wide approval (corporate identity). This part of management decisions may be described as "culture" management:

management of affective goals which influence the culture within a faculty in favour of willing to change education. Normally management decisions are focused on organizational structure variables (funding, roles and functions of departments).

Implementing Problem-Based learning in the Faculty of Policy and Administrative Sciences

The Faculty of Policy and Administration Sciences (FPAS) is the result of a fuse in 1988 between several established faculties, departments, institutions and programs concerning policy and administrative sciences. The creation of this faculty was a response to changes in societal need (society wanted graduates trained in various aspects of policy and administrative sciences) and to external pressures (a serious decline in student enrollment for some programs in policy sciences coupled with lower budgets).

The new program consists of four years divided into two cycles. The first cycle, year one and two, contains a broad introduction in policy and administrative sciences. These years are organized around courses with central themes from these sciences. They have a multidisciplinary orientation. The second cycle, the final two years, comprises courses with a more disciplinary character. The management of the program follows the structure of the curriculum. The first cycle is embedded and coordinated by the "school for policy and administrative sciences". The second cycle is organized by the individual departments of the faculty. The school concept is related with the idea of project teams ex article 40 according to the Dutch law on higher education. Daily coordination of the program is done by a program director and his staff. A council provides the director with complimentary advises and ensures the quality of the program.

The program in the first cycle is based on the principles of problem–based learning: multidisciplinary oriented courses and small group tutorials. The choice for problem–based learning as principal instructional method was a top–down decision. This decision was made in a period when FPAS had an interim management. The University of Nijmegen introduced interim management for FPAS because of earlier problems with the existing programs and problems risen as a result of the fuse. Teachers were compelled by the interim management to accept PBL as new teaching method. Teacher training was obligatory for staff members to learn the basic principles of PBL. This is, from a perspective of university organizations, an unusual and unconventional process of decision–making. However, strangely enough this didn't result in protests or faculty dissent. This may partly be explained by the fact that the process of merging resulted in a situation of general confusion for faculty members. Even more important was that a great number of staff members realized that holding their job positions was largely a result of increasing amounts of student enrollment due to the attractiveness of the new program.

The top–down approach of decision–making may partly explain why many staff members didn't feel responsible to transfer the PBL–principles to the particular needs of the FPAS program. Many teachers had a more positive attitude for the program in the second cycle (disciplinary oriented, traditional education) than for the first cycle (multidisciplinary and PBL oriented).

Conclusion

It is often thought that only new schools can implement PBL because of the radical changes needed in the organizational structure. The case described in this article proves that under certain conditions PBL can also be introduced in established faculties. In this particular case the explicit attention for management issues (interim management and introduction of the school concept) facilitated the introduction of PBL. Even more important are two other conditions. The first conditions is that external pressures were strong enough, in combination with changed management procedures, to resist internal forces wanting no change. The second condition is that some attention was paid to faculty approval and the culture of the organization. Management tried to change the organizational culture by explicitly focussing attention to compliance with the program, for example by training staff members in the principles of PBL at the University of Limburg. However, as such the management of FPAS has the feeling that when looking back into the past period much more attention should have been paid to cultural changes. More seminars, courses, workshop, internal magazines, festivities, financial stimuli to reward wanted organizational behaviour, should have been organized. Last but not least, student involvement and participation in the management of the program is needed to support the process of change.

References

Bess, J.L. (Ed.) (1984). College and University Organization. Insights from the behavioural sciences. New York: University Press.

Cohen, M.D., March, J.G., & Olsen, J.P. (1972). A garbage can model of Organizational Choice. Administrative Science Quarterly, 17, 1 - 25.

Goodlad, S. (Ed) (1984) Education for the Professions. Quis custodiet? (Guildford Surrey, SRHE & NFER-NELSON)

Hill, D.D. & Friedman, C.P. (1979). An analysis of frameworks for research on innovation and change in higher education. Review of Educational Research, 49, 411-435.

Kozma, R.B. (1985). A grounded theory of instructional innovation in higher education, Journal of Higher Education, 56, 300-319.

Lindquist, J. (1979). Strategies for change. Richardson, CA: Pacific Sounding Press.

Oldham, G.R., & Kulik, C.T. (1984). Motivation enhancement through work redesign.In: Bess, J.L. (Ed.). College and University Organization. New York: University Press.

Schmidt, H.G., & de Volder, M.L. (Eds) (1984) Tutorials in Problem-Based Learning.A New Direction in Teaching the Health Professions (Assen, Van Gorcum).

1247

INSTRUCTIONAL TECHNIQUES AND GROUPSIZE

Wim Jochems, Educational Development Unit, Delft University of Technology, Delft, The Netherlands

Introduction

In Dutch higher education a large part of the instruction-time available tends to be spent on lecturing, even in the case of rather small groups of students, which allow for more fruitful teaching procedures. In order to increase educational productivity, group size and instructional techniques have to be coupled more rationally in planning a curriculum. Therefore group size should depend on the interaction that is required in the process of mastering subject matter. From this point of view a problem-based and traditional curriculum are compared.

Problem-based and traditional curriculum in engineering education

Two years ago the Faculty of Building Sciences at Delft University of Technology started the implementation of a problem-based curriculum. The transition from a traditional curriculum to a problem-based one caused some dramatic changes because of a different planning and organization of the course.

Table 1 presents some data indicating the amount of varying educational activities programmed in a traditional curriculum and a problem-based curriculum, both in engineering education. The data are averages for the first en second year in the faculty of building sciences (problem-based) and in another seven engineering faculties (traditional curricula) at Delft University of Technology.

Table 1.

Programmed educational activities in hours per year and hours per week for a traditional and problem-based curriculum in engineering education (averages over first and second year).

curriculum	traditional		problem-based	
	per year	per week	per year	per week
lecture	530	19	110	3
instruction	130	4	430	12
lab work	300	11	250	7
total	960	34	790	22

In traditional curricula these activities are planned in four periods of seven weeks, each period being followed by examination weeks; a problem-based curriculum has six periods of six weeks each, including the examinations. Both types of curricula have re-examinations at the end of the summer holidays.

As can be seen in Table 1, there are remarkable differences between a problem-based curriculum and a traditional one. First, the average number of lectures per week is 19 hours in a traditional curriculum versus 3 in a problem-based. In general, these lectures can be characterised as monologues by a teacher providing explanations of the subject matter to large groups of students, on the average several hundreds of students.

Second, the average amount of instruction is 4 hours per week in a traditional curriculum versus 12 in a problem-based, one third of these being tutorials, also called tutor groups. Instruction refers to a classroomlike situation in which an instructor is setting tasks of various kinds to rather small groups of students, varying from about 10 to 40 at the maximum, providing them with cues, feedback, and so on. The kind of tasks presented to the students depends on the objectives of the instruction, which may be learning to learn, learning to design or learning to master problem solving routines, for instance mathematical ones.
If we don't take lab work into consideration because lab work isn't guided by teachers, three different teacher roles can be discriminated. The lecturer: the main task of the lecturer is to explain subject matter to the students. The tutor: his task is to support students in learning to learn; he teaches the students to handle problems in a systematic way, to formulate adequate learning objectives, to report orally in a brief and well-structured way; he also stimulates self-study activities (cf. Wilkerson & Hundert, 1991, pp. 163-167). The instructor: besides brief introductions he provides students with opportunities to practice certain skills, with cues and feedback and with (informal) tests to assess the student's mastery.

Educational productivity, group size and instructional techniques
What factors related to methods of instruction and learning processes are important for educational productivity? According to Guskey's study, 'students' involvement and active participation in instruction are extremely important to learning' just as 'providing students with regular and specific feedback' (Guskey, 1988, pp. 25-26). The educational productivity model as presented by Fraser et al. (1987, p. 205) shows that student participation, reinforcement, corrective feedback, and cues are important factors with respect to the method of instruction and the learning process.
Assuming these factors to be the most important now the question arises how to use the 'teacher-power' available in a curriculum. Leaving lab work out of consideration for reasons mentioned above, three types of teaching have to be programmed.
Lectures: explaining subject matter can be done in very large groups of students, e.g. hundreds of students; there is hardly a limitation to groupsize.
Instructions: active student participation with regular and specific feedback asks for classroomlike situations, e.g. 20 to 40 students to enable the instructor to control the learning process of each of the students individually, to give cues, to assess the progress and the results and so on.
Tutorials: if students have to work in close cooperation which requires a lot of interaction between students, e.g. in handling a case, the number of students should be reduced to 10 or 15 in order to enable them to participate actively.

Traditional and problem-based curriculum compared
What are the most obvious differences between a traditional and a problem-based curriculum? According to the data in Table 1, the amount of instruction (including tutorials) in a problem-based curriculum is about three times as large as in a traditional curriculum. Therefore, a problem-based curriculum offers more opportunities to activate students and to facilitate student participation. Whether these opportunities are utilized maximally, largely depends on the teachers.

Second, a problem-based curriculum asks for teachers who are able to perform different teacher roles. Depending on whether a lecture, an instruction or a tutorial is programmed, they should be able to use different sets of teaching techniques. Changing teacher roles has appeared to be very difficult (Driel, 1992). Third, a problem based curriculum is considerably more expensive than the traditional one. Nevertheless, the Faculty of Building Sciences could reduce costs to some extent by enlarging the number of students in instructions. Fourth, what about the learning outcomes in problem-based curriculum? Though some data are available concerning the Faculty of Building Sciences, favouring problem-based curriculum, it is very difficult to compare student results for the old and new curriculum, for instance because of changes in the content and the method of assessing learning outcomes.

In order to improve the productivity of the problem-based curriculum two measures have to be considered. First, the educational activities should be spread over a period of more than 36 weeks, just as in traditional curricula, to enable students to invest a greater amount of time in studying. Second, because teachers only partially utilize the opportunities offered by a problem-based curriculum, more attention has to be given to teacher training and faculty development.

References

Driel, J. van (1992). Faculty development as a tool in the implementation of educational innovation. Enschede: ECER-paper.

Fraser, B.J., Walberg, H.J., Welch, W.W. & Hattie, J.A. (1987). Syntheses of educational productivity research. International Journal of Educational Research, 11, 2, 145-252.

Guskey, T.R. (1988). Improving student learning in college classrooms. Springfield, Ill.: Charles C. Thomas Publ.

Wilkerson L.A. & Hundert, E.M. (1991). Becoming a problem-based tutor: increasing self-awareness through faculty development. In D. Boud & G. Feletti (Eds.), The challenge of problem-based learning. London: Kogan Page, 159-171.

MAGEMENT OF EDUCATIONAL CHANGE: A DISCUSSION OF THE IMPLEMENTATION PROCESS AT THE DEPARTMENT OF BUILDING SCIENCES

E. de Graaff, & J. van der Woord, Delft University of Technology, Delft, The Netherlands

Introduction

Management is a crucial factor in curriculum innovation. Fullan ironically quotes a deputy minister of education: "Well the hard work is done. We have the policy passed; now all you have to do is implement it", indicating that the real work only just begins when plans are accepted (Fullan, 1982, p. 54) The implementation of a new curriculum affects all members of the staff. The educators do not only have to learn new educational techniques, but also the traditional patterns of communication are uprooted. The old organizational structures often are unable to

cope with decisions that have to be made. Therefore, management and the design of new organizational structures are important aspects of the implementation process.

Strategies for change

Trying to implement PBL in an existing institution is asking for trouble (Schmidt, 1990). Moreover, the decision to opt for PBL at the department of Building Sciences was taken under external pressure (De Graaff & Van der Woord, 1990). The short time available and the strongly divided reactions to the external pressure at the department of Building Sciences made it impossible to attain general acceptance for any proposal.
The Faculty Board turned the external pressure into momentum for change, using it to speed up the process. As a strategy for change this can be characterized as a power–coercive approach (Chin & Benne, 1985). Several meetings were organized at the department to inform and coerce the members of the department. The Faculty Council however had to make a decision at short notice and without any real alternative. Actually at that time two different strategies were set in motion, one for the short term and one for the long term.

Getting started

At the short term all efforts were concentrated on persevering the momentum for change. Directly after the acceptation of the PKB–proposal (De Graaff & Bouhuijs, 1992) by the council a series of training sessions and workshops with the prospective members of the first blockgroups was organized. This way a group of supporters for the new curriculum was formed. Next a new organization structure was build beside the existing one. The Faculty Board installed a committee called Implementation Committee Educational programme Building Sciences (ICOB), charged with the responsibility for the development of the new curriculum. Members of ICOB were selected on personal title and not as representatives of any other body within the department. Downward from ICOB a linking–pin structure was designed. Each ICOB–member was assigned a part of the curriculum (the first year, the second year and the graduation majors), chairing the respective curriculum–groups, consisting of the block–coordinators, which in turn chair their block–groups.

A long term strategy for curriculum–innovation

The short term strategy described above has some severe drawbacks. The initiative rests with a small and rather isolated group. This serves well to get the train going, but in order to sustain motion, broad support among the staff is necessary. The method of PBL has been developed to suit the needs of medical education in McMasters and in Maastricht (Boud & Feletti, 1991). It is easy to predict that people who feel they have not been let in on the decision will reject this method as "not invented here" and "not suitable for our needs".
In a sense these people are absolutely right. The introduction of PBL at the department of Building Sciences is not really an innovation but rather the application of an improvement (De Zeeuw, 1990). And even as an improvement the PBL–method from Maastricht (or anywhere else on the world) is bound to fall short if it is applied straightforward in a completely different situation. The success

of curriculum–innovation in the long run depends on the ability of the department to adapt the educational method to suit its own specific needs. In order to be truly innovative the new curriculum should result in increased competence for the organization and its members.
An important advantage of PBL is the introduction of a new way of thinking about education and learning. In the old situation there was little reflection on the task of teaching. Discussion on education was mostly confined within the disciplinary boundaries. The PBL–curriculum breaks down these traditional boundaries. Discussions on the educational approach are further encouraged by the availability of detailed educational programmes (the blockbooks). At a conference in February 1992 the Faculty Board explicitly emphasized the importance of continuous improvement of the educational method in the course of adaptation to the needs of Building Sciences. The installation of an educational unit within the department is an important factor in the strategy of on-going innovation, embodying the systematic approach to educational development, programme evaluation and faculty development.

Human Resources Management
The capital of a university is mostly invested in the expertise of the teaching staff. Since the essence of educational change is learning new ways of thinking and doing, new skills, knowledge, attitudes, etc., it follows that staff development is one of the most important factors related to change in practice (Fullan, 1982, p.66). The problem is that it can be very difficult to teach old dogs new tricks. University teachers with twenty odd years of experience have some right to say "I've done this job for so long without criticism, what's the point in changing now?"
The department, however, has to deal with changes in the external situation (De Graaff & Bouhuijs, 1992). In order to meet these demands some form of control over the availability of expertise is necessary. Within the context of Human Resources Management (HRM) a strategy has been developed based on selection, assessment and tailored trainingprogrammes (Vloeberghs, 1989). The aim of this strategy is to gradually reevaluate the staff and provide opportunities for further personal development in order to attain a more flexible educational organization. Together with the strategy of on-going innovation this should result in increased competence for the institution as well as its members (individual and organizational learning).

References
Chin, R. & Benne, K.D. (1985). General strategies for effecting changes in human systems. In W.G. Bennis, K.D. Benne & R. Chin, The planning of change (fourth edition). New York: Holt, Rinehart & Winston.
De Graaff, E. & Van der Woord, J. (1990). Changing horses mid–course. Paper presented at the Second International Symposium on Problem–Based Learning, Yogyakarta. Indonesia October 7–12–1990.
De Graaf. E & Bouhuijs, P.A.J. (1992). The introduction of Problem–based learning at the department of building Sciences. Enschede: ECER–paper.
Fullan, M. (1982). The Meaning of Educational Change. New York/London: Teachers College, Columbia University.

Schmidt, H.G. (1990). Onderwijskundige aspekten van probleemgestuurd onderwijs. In W.S. Jochems (red.), Activerend Onderwijs, over onderwijsvormen die het leren bevorderen [Activating Education, on educational formats that encourage learning]. Delft: Delft Universitaire pers.

Vloeberghs, D. (1989). Human Resources Management. Leuven/Amersfoort: Acco.

De Zeeuw. G. (1990). Problemen van verbeteren en innoveren [Problems of improvement and innovation]. In Postdoctorale opleiding Innovatie manager. Amsterdam: Academie voor Informatica, Universiteit van Amsterdam.

CHAPTER 9

THE SOCIAL CONTEXT OF EDUCATION

THE SOCIAL CONTEXT OF EDUCATION

Roel J. Bosker, University of Twente, Enschede; & Anton Wesselingh, University of Nijmegen, Nijmegen, The Nertherlands

Introduction

What are the social determinants and effects of education? These are the central questions of sociology of education. As a seperate displine within educational science, the key concept is "social". With this concept reference is made to groups or to society at large instead of to individuals. Education is viewed as the arena in which social groups fight for their privileges and power. But it is also viewed as a key mechanism to promote individual or group mobility as well as social change. The European Conference on Educational Research highlights important contributions to our understanding of several central topics. The most important contribution being that of relativism: as contributions from about 15 countries will be presented social and cultural factors that vary over these countries will appear to have their own substantial effects. Although a comparative point of view will not be the topic of the papers to be presented, the audience will be confronted by the cross-national differences, thus being provided with alternative views on stratification and education.

An important aspect of many papers is the concept of change, whether it be for national educational reform, the restructuring of the society at large (as is the case in eastern-european countries), large scale innovations to improve the educational system, or small scale interventions: answers are sought to the question whether these changes really should be valued as improvements in terms of increased openess of the society, i.e. more individual and group mobility. Next to this an important part of the papers tries to unravel social, cultural and even financial resources in educational attainment, thus leading to a better understanding of the determinants of inequality in educational and status attainment. What are these resources? Generally speaking cultural resources refer to the cultural capital of the primary and secondary socializing agents, being the family in which the child is raised respectively peers, teachers, public media, etcetera. This cultural capital (the concept stems from Bourdieu, 1984), the relevance of which is being assessed primarily by its correspondence to the culture transmitted in school, manifests itself in the pedagogical interaction between the child and its socializers. Important in this respect are the origin and determinants of lifestyles of young people, as it will be addressed in an Anglo-Saxon paper. The social capital of the family (this concept is recently being stressed by Coleman; e.g. Coleman & Hoffer,

1987) refers primarily to the social networks available to the family: one-parent families, low status families, ethnic minority families, etcetera, have a more confined and restricted network than other families, thus leading to less opportunities for the children raised in these families. Financial resources, although in previous times seen as the primary source of inequality in education, have now lost most of their significance in western countries, although one of the papers will demonstrate that these resources still can explain part of the inequality in higher education.

A further point of discussion, originally brought up by Boudon (1974) has to do with the question whether inequality in education is mainly being caused by the primary or the secondary effects of stratification: i.e. by the direct effects of stratification on achievement versus the effects of stratification on choices that individuals have to make at the various branching points in their school career (in differentiated school systems this is the track in which the pupil is placed; other choices refer to dropping-out or staying-in, choosing vocational or general education, proceeding with education on the higher level or entering into the labour market, and so on). The papers focus on these decisive points, and the substantive contributions try to make clear whether rational choices of individuals are the only causal explanation of these phenomena. It turns out that not only individual choices, but also latent yet powerful sources of selection in the system are at work. These are not based on achievement criteria only but on social characteristics as well (such as they manifest themselves in teacher behaviour and expantancies). Interestingly enough papers will be presented that show that these processes occur as early as in primary education as well as in the transition from education to the labour market. In this latter case rational choice theory has been predominantly manifest in human capital theory, stating that choices of individuals can be understood as choices to invest in education in order to improve the later labour market position. Screening theory, more or less competing with this economic approach, will be used in one of the papers, stating that employers have different selection criteria in mind as perseverance, creativity, etcetera. It is also interesting that some of the papers will adress the question whether or not macro-societal constraints (legislation, second chance education) are active in these issues.

Very important are those contributions that more or less evaluate interventions to bring about more equality in educational opportunities for girls, ethic minority children, and low status pupils. These interventions start very early, in order to improve the starting positions of deprivated children at school entrance. Interestingly enough interventions go beyond the direct boundaries of education as some of these interventions seek to bring about changes in the social and cultural capital of the parents. A symposium will be held as part of the theme on a more or less holistic intervention (school, family, neighborhood) to improve educational opportunities of ethnic minority children in The Netherlands. Other interventions specifically seek to improve the educational opportunities of girls, with emphasis on girls and mathematics, girls and information technology, and woman and 'second chance education'. Moreover, a research review on this specific topic, encomprising several decades of research on gender and education, will be presented at the conference.

Despite the great variety of topics covered by the papers cross-national issues are not addressed yet, whereas the integration of Europe, for instance, raises new important questions on equity, such as caused by the problem of certification and the subsequent opportunities on the labour market. Cultural pluralism in education is one of the topics that may become of growing interest in rapidly diversifying societies. In this respect the key-note address of Prof.Dr. Heiner Meuleman will be of special importance, since it seeks to answer some of the topics of cross-national interests by cross-national comparisons, a.o. the topic of educational expansion and equity.

Although highlighted in terms of contributions to the knowledge base of sociology of education the substantive topics of the various papers can be summarized in four broad categories:

- the study of educational opportunities for ethnic minorities;
- the study of educational opportunities for women;
- the transition from education to the labour market;
- miscelleanous topics, as, for instance, family factors in achievement, legislation and equity, vocational and higher education, etcetera.

References

Boudon, R. (1974). Education, Opportunity, and Social Inequality. New York.

Bourdieu, P. (1984). Distinction, a Social Critique of the Judgement of Taste. London.

Coleman, J.S. & Hoffer, T. (1987). Public and Private High Schools, The Impact of Communities. New York.

1257

GENDER AND EDUCATION: AN EXPLORATION OF RESEARCH PROGRAMMES

Edith van Eck, Geert ten Dam, & Monique Volman, Centre for Educational Research, Graduate School of Teaching and Learning, University of Amsterdam, Amsterdam, The Netherlands.

Introduction

At the end of the seventies, the first studies on sex inequality in education in the Netherlands were published. This paper is based on a survey of Dutch emancipation research and women's studies in the field of education (Ten Dam, Van Eck & Volman, 1992). In this study we analyzed the starting points, methods and research questions as well as the results of research in this field. The purpose of the present study was to describe developments in this research area during the past ten years relating them to the research-external and research-internal context and to explore blind spots.

Methodology

In the study reported here the research area was analyzed by means of the 'reconstruction of research programmes' (Hetebrij, 1983; Hetebrij & Nijssen, 1986). This method focuses on the extent of coherent clusters of theoretical and empirical studies in a particular field, and tries to identify the theoretical perspectives connecting the different studies and determining the kind of questions asked, methods used and the resulting accumulation of knowledge.

We gathered all research literature on gender and education, both reports and articles, that has been published in the Netherlands since 1979. This literature was described and analyzed on the basis of a questionnaire concerning the research tradition, theoretical framework, methods, research questions, the institutions and bodies for which the research was carried out, the results and the discussions with other researchers on theoretical framework, methods or results.

Firstly, the study produced a survey of the research on gender and education and the results of this research in relation to the Dutch context. We described the national policy on equal opportunities in education and the financing of educational research in the Netherlands. Secondly, we traced the occurrence of research programmes in the projects described. Finally we presented an overview of developments in Anglo-Saxon research on gender and education based on a number of reviews and several interviews with foreign experts.

1258

We classified the studies described as follows:

- educational choices and educational careers;
- evaluation of the educational policy concerning equal opportunities;
- immigrant girls in education;
- higher education;
- subjects and courses traditionally chosen by women;
- adult education;
- careers of women teachers;
- science and technology ;
- learning, teaching and classroom interaction.

Results and conclusions

Our analysis of research projects in this area shows an increasing level of specification of questions asked and answers produced. Furthermore, we found a shift in theoretical framework. In the first period an analysis in terms of 'reproduction of inequality' and the role of educational institutes in selecting, socializing and allocating individuals to socially suitable positions predominated. Later on the concept of effective schools emerges and the meaning of effectivity is no longer a topic of discussion; effectivity is almost exclusively related to educational achievement. We also noticed a shift in the themes under investigation, sometimes related to a shift in educational policy, a phenomenon one can expect in policy-oriented research. More and more accent is put on the role of education in qualifying for the labour market in general, but also for special groups, for example, women or immigrants. Education as a tool for personal development gradually looses ground in the period under investigation. Finally, the greater part of the research projects is aimed at exploring and describing correlations between variables, research on underlying processes and explanations being relatively underrepresented.

Even in policy-oriented research, where research questions arise primarily from educational policy and not from educational theory, several research programmes appeared to exist. This was especially the case when research on a theme had been subsidized for a longer period. A situation where researchers on gender and education can become affiliated to existing theoretical traditions referring to broader research areas, (e.g. research on effective schools) and can use existing data and instruments, promotes the development of research programmes. It was also apparent that once a research programme has been started researchers reformulate the research questions in terms of their theoretical framework rather than following the terminology of the institution or body financing the project. The theoretical development is thereby promoted by the incorporation of research into existing programmes.

The research programmes we identified are not fully developed yet. Furthermore, none of them was specifically developed for research on gender and education. Finally, we found a lack of discussion between researchers representing different theoretical frameworks in this field. The discussions identified were primarily about methods and research results, not the theoretical starting points.

References

Acker, S. (1990). Gender issues in schooling. In N. Entwistle (ed.), Handbook of educational ideas. (p.91-99). London: Routledge.

Dam, G. ten, & Volman, M. (1991). Conceptualizing gender differences in educational research: the case of the Netherlands. British Journal of Sociology of Education, 12 (3), 309-321.

Dam, G. ten, Eck, E. van, & Volman, M. (1992). Onderwijs en sekse: een verkenning van researchprogramma's. Vierde trendrapport STEO. Den Haag: Directie Coördinatie Emancipatiebeleid/ Stimuleringsgroep Emancipatie-Onderzoek.

Eccles, J.S. (1987). Gender roles and women's achievement-related decisions. Psychology of Women Quarterly, 2, 135-172.

Eck, E. van, & Volman, M. (red.) (1990). Determinanten van seksespecifieke keuzen en prestaties. Verslag van symposium gehouden op de Onderwijsresearchdagen 1989. Amsterdam: SPCP.

Harbers, H. (1986). Sociale wetenschappen en hun speelruimte. Groningen: Wolters-Noordhoff.

Hetebrij, M. (1983). De ontwikkeling van researchprogramma's als methodisch probleem. Amsterdam: VU Boekhandel/Uitgeverij.

Hetebrij, M., & Nijssen, A.J. (1986). De reconstructie van onderzoeksprogramma's: een instrument voor inhoudelijke programmering. In M. Hetebrij & A.J. Nijssen (red.), Researchprogramma's en beleidsonderzoek, een primair inhoudelijk probleem (p.15-38). Amsterdam: SISWO.

Kuyper, H., m.m.v. Snippe, J., & Werf, M.P.C. van der (1990). Probleemverkenning Emancipatie *1990*. 's-Gravenhage: SVO.

Ministerie van Onderwijs en Wetenschappen (1985). Derde Nota onderwijsemancipatie: een stukje wijzer. Den Haag: Ministerie van Onderwijs en Wetenschappen.

Stimuleringsgroep Emancipatie-Onderzoek (1988). STEO-advies aan het Instituut voor Onderzoek van het Onderwijs (SVO) en de bij onderzoeksprogrammering van SVO betrokkenen. 's-Gravenhage: Stimuleringsgroep Emancipatie-Onderzoek.

ARE GIRLS DISADVANTAGED IN A COMPUTER LITERATE SOCIETY?

Jean Underwood, Susan Cavendish, School of Education, University of Leicester; & Geoffrey Underwood, Department of Psychology, University of Nottingham, United Kingdom

Abstract

Schoolchildren were asked about their home and school technological environments, and about their competence with technology. Girls reported that they live in technologically less rich environments than boys, but our initial premise, that the children's depth of experience with home technologies would influence the development of IT skills in the classroom, proved too simplistic. The relationship between IT attitudes and skills acquisition and individual differences in age, gender and experience proved to be complex .

Introduction

This study is part of an investigation into the influence of the richness of children's home technological environment on the use of Information Technology (IT) in the classroom. The implementation of a National Curriculum for England and Wales has moved IT from the periphery to the heart of the UK curriculum for all students. The focus of the study is on the ability of students to benefit from the new curriculum.

There is a considerable research literature linking home factors to the development of academic skills such as reading. Should we expect similar influences of the home technologies on children's IT skills? The experience which children have of home technologies may provide important training for the development of IT skills and understanding, for many tasks such as telephoning Granny or using a VCR to record a favourite television programme are meaningful, planned actions for children.

The relationship between home and school learning is not simple, however. Resnick (1987) argues that *"There is growing evidence, then, that not only may schooling not contribute in a direct and obvious way to performance outside school but also that knowledge acquired outside school is not always used to support learning inside school."* Greenfield (1984), in discussing the transfer of skills from game playing to classroom IT, suggests that such transfers may occur at only the simplest level, if at all, because children are unlikely to articulate the knowledge they gain through such game playing: *"transfer of concepts to a new domain often seems to require their formal verbalisation: yet the knowledge*

1261

gained in playing video games is more than likely non-verbal." In asking the question 'Do experiences of home technology affect classroom IT?' we need to be aware that the degree of transfer that we find might be limited indeed!
Studies of the influence of prior computer experience on IT skills acquisition provide contradictory results although the overall evidence shows that the quality of that experience has profound effects on relieving or exacerbating computer anxiety. In the latter cases, where computer anxiety increased, the novice courses were based specifically upon computer programming. There is also evidence that the beneficial effects of computer experience are constrained by gender, although opposing accounts are presented. Todman and Lawrenson (1992) have shown that there is an interaction between age, computer anxiety and computer experience. Their finding that 9 year olds were less computer anxious than 18 year olds was attributed to the greater and earlier computer experience of the younger subjects coupled with an introduction to computers in a non-threatening environment both at school and home.
The question arises as to who is likely to have had this beneficial prior experience. Home ownership of computers has been identified as being of particular importance in forming attitudes to computers and the subsequent development of machine competence, and Culley (1988) reported that home computers are six times more likely to be bought for boys than girls. Do girls, therefore, experience an impoverished technological environment and subsequently suffer disadvantage in their acquisition of IT skills? We conducted a survey to investigate this question.

Method

Subjects: 201 pupils aged between 7 and 12 years of age took part. There were 100 boys and 101 girls. The pupils attended one of four schools situated in mixed social class areas. Two schools were primary schools with pupils from 4+ to 11 years, and two schools were high schools with pupils from 11 to 14 years of age.

Instruments

Subjects were presented with a questionnaire to determine the richness of their home technological environment; their personal ownership of a range of technologies; and their use of computers. The data collected were as follows:

- pertinent background information on sibling position and gender ;
- overall richness of the home technological environment;
- personal ownership of technological devices;
- self-reported and measured competence with technological devices;
- experience with and attitudes to computers both in the home and at school.

Procedure

In the primary schools, children were grouped into four and taken to one side of their usual classroom where the experimenter presented the questionnaire. This procedure was adopted in order to minimise test anxiety and to ensure that children understood the questions. In the high schools, the questionnaire was administered in a class situation. There was no time limit for completion.

Analysis

Frequencies of responses were obtained for all questionnaire items and Chi-squared tests applied to compare responses by gender and by age of the children. Four composite variables were constructed for further analysis: *richness of the home technological environment; gadget ownership;* the child's *self-reported competence with technology;* and *measured competence with technology.* The score for each child's *home technological environment* was calculated by awarding a point for every gadget that the individual had in the home (maximum score 15) as was *gadget ownership* (max.10). *Self-reported competence with technology* was a count of the number of specified operations the child reported he or she could accomplish (max. 22). *Measured competence with technology*, the child's ability to produce a list of sequenced instructions, was scored by awarding a point for each correct instruction that the individual noted down (max. 32).

Results

The results show that both boys and girls are living in rich technological environments with the majority of children claiming home ownership of everyday items such as telephones, washing machines, video recorders and cassette players (Mean=11.6 out of a maximum of 15 points available; s.d=2.0). Microwave ovens and dishwashers are found in about half of all households. These figures follow a similar pattern to official statistics, except in the case of high incidence of computer occurrence.

Strong positive relationships were found between richness of the technological environment, ownership of technology, self-reported competence, measured competence in the use of technologies, and the range of computer activities engaged in at home and school (Table 1).

	GO	*SC*	*MC*	*RUH*	*RUS*
Tech. Environment (TE)	+.52	+.47	+.21	+.27	+.13
Gadget Ownership (GO)		+.40	+.18	+.31	+.03
S-Report Competence (SC)			+.42	+.48	+.19
Measured Competence (MC)				+.21	+.08
Range of Use at Home (RUH)					+.21
Range of Use at School (RUS)					

Table 1

Pearson Product-Moment Correlations. Correlations between the measures of technological environment and technological competence. Figures in bold are reliable at p<0.01 or better; N=201.

The analysis of home and personal ownership of technologies by gender revealed that boys are not only more likely to live in a home where there is a computer (c^2=6.5, df=1, p <0.01) but they are also more likely to be the owner of that computer (c^2=12.3, df=1, p <0.001) than are girls. These figures support those reported by Culley (1988) and by Underwood and Underwood (1992). In addition

1263

boys are also more likely to own a video game, a robotic toy or a digital watch than girls, while girls own the greater proportion of analogue watches and typewriters. The incidence of an older brother also increased the likelihood of a home computer being present (c^2=7.0, df=1, p <0.01). These results were largely unaffected by age. The youngest children (7 and 8 year olds) did report fewer gadgets in the home, but the children's memories of their environments may be responsible for this difference in their reporting of technological richness, of course.

Additional analyses of the use of computers showed that boys and girls performed similar tasks within schools but their out-of-school usage was very different. Girls used the computer to write stories and play mathematical programs. Although about half the boys wrote stories with their computers they were also using a wide range of application packages to draw and play music, for information handling and programming and they were more likely to be involved in computer 'housework' such as disc copying.

There was a significant relationship between positive attitude to computer use and the range of activities in which children engaged. This relationship held for both home (c^2=73.6, df=1, p<0.001) and school (c^2=128.1, df=1, p<0.001) computer experience but experiences at home did not lead to positive attitudes at school (c^2 =10.4 df=1, *n.s.*) and there was similarly no transfer of regard from school to home computers (c^2=17.1, df=1, *n.s.*)

Discussion

In summary, the relationships between richness of the technological environment, ownership of technology, self-reported competence and measured competence in the use of technologies show that the greater the number of gadgets in the home the greater the competence with technologies in general.

Although the majority of pupils live in technologically rich environments there are significant differences in the ownership of, and interaction with, the technologies for boys and girls. Boys are more likely than girls to own and operate technologies with direct input to school IT experiences, such as robotic toys, computers and programmable digital watches. At home boys use a wide range of computer packages but girls are more restricted in their exploration of the use of computers. Positive attitudes to computers are related to the range of computer use but that experience is context bound. Wide home computer use does not necessarily lead to positive attitudes to school IT and increased use of IT in schools does not necessarily generate an enthusiasm for home computers.

This study has shown that the relationship between IT attitudes and skills acquisition and individual differences in age, gender and experience is complex. This confirms the findings of Todman and Lawrenson (1992). We pose three key questions from our research:

- Are girls disadvantaged in the IT curriculum because of their lower home exposure to technology? There is some evidence that successful classroom experiences can, in part at least, compensate the girls for the earlier out-of-school computer experiences of boys (Underwood and Underwood, 1990);
- Why is it that home computer literate children, who predominantly are boys, do not transfer positive attitudes to the classroom?

- Are girls choosing to be technologically disadvantaged or are factors such as parental and societal expectations inhibiting their technological education?

References

Culley, L. Girls, boys and computers. Educational Studies, 14, 3-8, 1988.

Greenfield, P., Mind and Media. Harvard: Harvard U.P.,1984.

Resnick, L. Learning in school and out. Educational Researcher, 16,13-20, 1987.

Todman, J. and Lawrenson, H. Computer anxiety in primary schoolchildren and university students. British Educational Research Journal, 18, 63-72, 1992.

Underwood, J. & Underwood, G., Computers and Learning: Helping Children To Acquire Thinking Skills. Blackwells: Oxford, 1990.

Underwood, J. & Underwood, G., Predicting computer literacy. In the Proceedings of the 9th International Conference on Technology and Education, 2. Austin: University of Texas, 1992.

1265

DETERMINANTS OF SEX-SPECIFIC SUBJECT CHOICES IN SECONDAIRY EDUCATION

H. Dekkers, Instituut voor Toegepaste Sociale Wetenschappen (ITS), Nijmegen, The Netherlands

Subject choice in secondary education is still different for boys and girls. Especially mathematics and physics are subjects that are not chosen by female pupils. These subjects however are very important for admission to and success in future studies and the labourmarket.

We assume that in the first years of secondary school a process takes place, in which differences between boys and girls increase and mutually change. That's why we try to investigate in a longitudinal study what these differences are and whether there are changes in the effect that crucial variables have on the subject-choices of boys and girls in AVO/VWO.

For the above purpose a group of 2500 pupils from 25 schools is followed from the first year of secondary education (brugklas) until the third or fourth year, when they have to make their choices. Different kind of variables, all derived from previous research in this area, have already been measured for years. Every year the pupils make a CITO-test for mathematics and fill out a questionnaire in which their background and attitudes are measured. Summarized the variables concern: parental background, previous educational results, perspective of the future, attitudes towards subjects (a.o. attribution of success and failure) and influence of relevant others.

We can already report some results. In June 1992 all the results of the first two years will be available; for the definite paper a selection will be made.

Some tentative results are:

- the first year girls and boys differ in CITO-test results; the second year the differences are not relevant;
- the CITO-test results differ between schooltypes and (less logical) also between schools;
- in the second year there are sex-differences in the choice of the subjects physics, French and MAVO/HAVO mathematics, not of VWO mathematics a and b;
- the correlation between schoolmarks and subject choice is bigger than between CITO-results and subject choice, and the correlation between marks and CITO is smaller than expected. These correlations are not sex-specific, but on a different level for boys and girls;
- the interst in further studies is still very sex-specific and traditional, and related to subject choice;

- the plans for future family-life and work are very sex-specific, but not related to subject choice. The plans of girls and boys are fully complementary in this (st)age;
- girls consider mathematics less necessary and suitable for their future, boys French; a positive judgement is related to subject choice;
- attribution-items are less correlated with subject choice than judgement variables;
- judgements about, and choices for physics and French are often reversible;
- parents are the most important and most sexist advisers.

A combination of some of the above results will be elaborated in the definite paper. Not only sex-effects, but also school(type)-effects will be subject of discussion.

1267

MALE AND FEMALE STUDENTS AT THE OPEN UNIVERSITY OF THE NETHERLANDS: LABOUR-MARKET PARTICIPATION AND ACCESS TO RESOURCES

José Janssen, & Gerry Joosten, Department for Research and Evaluation, Open University of the Netherlands, Heerlen, The Netherlands

The foundation of the Open university of the Netherlands was accompanied by discussions on the number of students enrolling from different target–groups.
Target figures for women's participation in the Open university student population varied from about 40% up to 60%.
Estimates based on the proportion of women in the Dutch population expect a percentage of female students of at least 50%. We might call this the 'reasonable' view. The 'realistic view' takes into account that women have a disadvantage in secondary education to be caught up first and expect a percentage of about 40% to 45% of female students.
Actually, the proportion of female students at the Open University of the Netherlands has reached what seems a stable state of about 38%.
In order to increase our understanding of the participation rates of female students we should direct our attention to the question which categories of women participate.
Comparing the total student population with the Dutch population in respect to labour-market participation, it appears that Open university students are relatively often employed. This is especially the case with female students: 70% of the female students is employed, compared to about 50% of the Dutch women. Besides, they differ in yet another respect. Compared to Dutch women, female students are not only employed more often, but also hold full-time jobs more frequently: 64% of these students work at least 35 hours a week, compared to 45% of the employed Dutch females.
From these figures we might conclude that the Open university attracts a specific category of women, resembling the male population at the point of labour-market participation.
Explanations of higher participation rates for employed people in the field of adult education, often are linked up with human capital theory. Education then is conceived of as an investment in ones skills. The costs of the acquisition of these skills are expected to be compensated by future returns. The decision to take up

study again, is based on a process of weighing costs against future payoff rationally.
On the basis of human capital theory one can argue that employed people are more likely to participate in adult education, because an investment in education bears less risks for them: the future payoff is more certain.
Moreover employed people generally are more resourceful. They earn a salary and sometimes receive incentives from the employer like paid leave and compensation for costs of study.
One of the criticisms towards human capital theory is directed towards the fact that it does not take into account the unequal access to resources. In this context we investigate some resources employed male and female students have access to. Though ideally, the comparison should be made between employed people who do take up study and employed people who don't, an investigation of the resources employed male and female students have access to, can give us some clues on the validity of human capital theory.
Our analysis is based on data collected from the student population of the Open university in the Annual Educational Research of 1991 and on data from the enrolment forms of the last course students took to date.
A sample of 2020 students was drawn from the total student population of 33,533. Data were collected through a postal survey.
The response rate was 61% which is quite satisfactory taking into account that about 30% of the students starting a course at the Open university do not complete the course. A weight procedure was carried out in order to compensate for the under representation of these students. The weighted response is representative for the whole student population.
The resources that will be considered in our analysis are paid leave and compensation for costs of study. Unfortunately we won't be able to relate the use of these resources to income, because we don't have any data on students earnings. Since we are dealing with nominal variables the relationship between the use of resources and sex is explored by analyzing cross tables. Chi-square tests will be used to examine the significance of sex-differences.
Employed male students receive compensation from employers more often than female students (38% compared to 25%). In order to draw any sound conclusion on differences between male and female students concerning use of resources, we should control for the relevance of the study for the present job.
Employed male students more often reported that the study is relevant for their present job than female students (44% and 26% respectively). Controlling for this reported job relevance, there is no significant difference between male and female students in respect to compensation for costs of study or paid leave.
The difference in reported job relevance however is remarkable since there is no similar difference in study motives mentioned on the enrolment forms: 28% of the employed male students mentioned 'improving performance in present job' as a motive to take up study, compared to 21% of the employed female students. This difference in study motives does not match the difference in reported job relevance.
These figures suggest that male students interpreted 'relevance' in a broader sense than female students. It might therefore be a mistake to rely on students reports of job relevance.

Using the study motive 'improving performance in present job' as a second indicator for job relevance (the first being the reported job relevance from the questionnaire), the results are quite different. In this case our conclusion would be that there is a significant difference between employed male and female students concerning use of resources, due to the fact that male students more often receive compensation of costs, apart from the relevance of the study for the present job.
Interpreting the results based on *reported* job relevance; does not have any consequences for the theory: the use of resources differs for both sexes but this difference does not originate from unequal access, but from a different context of taking up study: study activities of female students are less often directed towards the present job. In terms of human capital theory: female students seem to expect the returns for the investment from another job more often and therefore do not equally come in for compensation for costs from the present employer.
On the other hand, relying on the *study motives* mentioned on the enrolment form as an indicator of the relevance of the study for the present job, the results do affect theory: the use of resources differs for both sexes, because male students more often receive compensation for costs even though the relevance of the study for the present job does not appear from their study motives.
Drawing upon other research we could explain this difference between employed male and female students from gender related differences in attitudes towards ones job: male employees tend to stress the rights associated with their job more often than female employees. This indicates that the weighing of costs and returns differs between both sexes, along with different perceptions and interpretations of the environment.

1270

SOCIAL INEQUALITY AND THE DEMAND FOR POST-COMPULSORY EDUCATION: A MICRO-ECONOMIC APPROACH

Ides Nicaise, Hoger Instituut van de Arbeid (HIVA), Catholic University of Leuven, Leuven, Belgium

In this article, the problem of the democratization of education is approached from a micro-economic point of view. The "socially blind" standard human-capital model is transformed by taking into account imperfections in the capital market, the "consumption aspect" of education, and the complementarity between school and family inputs in the "production" of human capital. An empirically testable demand function for education is derived, in which "full income" (including the shadow income of parental homework), direct and indirect study costs (see table 1) and family background variables enter as arguments. The model is fitted by means of logistic regression analysis on a cross-section data base for Flanders (1985).

Table 1
Composition of the average study cost (in BEF) in upper-secondary and higher education (1985-86).

	USE	HE
Direct cost	15,600	58,600
Indirect cost	315,900	335,800
Total gross cost	331,500	394,300
Subsidy	74,500	89,100
Net cost	257,000	305,200

A first variant of the model (see table 2) examines the probabilities of participation as such. It yields the following conclusions:

- the participation in higher education is clearly income sensitive, which does not seem to be the case for upper-secondary education. In higher education, each increase of the full income by 1% causes the educational participation to increase by 0.6 percentage points;
- the estimations reveal a strong degree of cost sensitivity in the demand for education (cost is interpreted here as the sum of direct expenses linked to education and foregone earnings of students). Moreover, our results are in line with those of Fuller et al. (1983) in the sense that the cost sensitivity declines with increasing income. Logically, it follows that financial compensations to families

1271

with studying children must be efficient too, and that their efficiency increases with their selectivity;

- Family size appears to have two opposite effects on the demand for post-compulsory education : a negative direct effect, through a reduction of the disposable income per capita; and a positive indirect effect, through the higher average social transfers granted to students from large families. These transfers seem to (over)compensate the direct negative influence of family size on participation probabilities;
- Further, we note that the educational level of the mother also has a clear positive effect on the participation of the children. Youngsters whose mothers have an upper-secondary education certificate (MOTH_USE=1) show a 32% higher probability of participation in upper-secondary education than the others; for higher education, the participation probability rises by 36% if the mother has a diploma of that level (MOTH_HE=1);
- Finally, gender has a significant impact in higher education, no longer on first-year enrollment but on overall participation : boys still participate more in higher education than girls do. The difference between first-year and overall enrollment gaps can be attributed partly to higher drop-out rates for girls, partly to their orientation to shorter studies, and also to the consequences of earlier inequalities in enrollment.

Table 2

Estimation results of the basic model

UPPER-SECONDARY EDUCATION:				
Chi square of the model: 83.4***		C test: .80		
variable	beta	Std.dev.	chi sq.	R
INTERCEPT	14.91*	7.98	3.5	-
LOG(FULL INC.)	- .14	1.39	.01	.00
LOG(FAM.SIZE)	.26	.60	.19	.00
LCOST(LOW.INC)	-4.92***	1.37	12.86	-.17
LCOST(MID.INC)	-4.92***	1.35	13.23	-.17
LCOST(HIGH.INC)	-4.63***	1.34	11.88	-.16
MOTH_USE	1.24***	.43	8.33	.13
MOTH_HE	7.23	22.05	.11	.00
MALE	.59	.46	1.61	.00
HIGHER EDUCATION:				
Chi square of the model: 90.6***		C test: .80		
variable	beta	Std.dev.	chi sq.	R
INTERCEPT	10.75	7.26	2.19	-
LOG(FULL INC.)	2.82**	1.32	4.51	.08
LOG(FAM.SIZE)	-1.71***	.65	6.97	-.11
LCOST(LOW.INC)	-7.46***	1.42	27.50	-.25
LCOST(MID.INC)	-7.33***	1.39	27.94	-.25
LCOST(HIGH.INC)	-7.13***	1.36	27.43	-.25
MOTH_USE	.22	.32	.47	.00
MOTH_HE	1.83***	.65	8.02	.12
MALE	.93***	.34	7.26	.11

*: level of significance = 90%; **: 95%; ***: 99%

A refined version of the demand model distinguishes between 3 or more options, including non-participation besides different types of study orientations : general (GSE), technical (TSE) and vocational (VSE) education at the secondary level; university and non-university orientations at the higher education level).

- The estimates regarding upper-secondary education yield little, just that the demand for general education is significantly less cost sensitive than that for technical and vocational education. Further, it appears that boys opt more for technical education than do girls;
- As regards higher education, the estimates point to a greater *cost* sensitivity of the demand for higher non-univesity education (HNUE). On the other hand, non-university education is less *income* sensitive than the university;
- Finally, the influence of sex on the choice of study is striking: the university is still more attractive for men while the (shorter) non-university orientations attract more women. This confirms one of the explanations of the lower overall female participation given above.

Table 3

Results of the multiple choice model

UPPER-SECONDARY EDUCATION:			
Chi square of the model:	565.71***	N = 269	
effect on	GSE P_1/P_N	TSE P_2/P_N	VSE P_3/P_N
INTERCEPT	5.47	16.03	17.47*
LOG(FULL INC.)	1.09	- .01	- .50
LOG(FAM.SIZE)	.21	.04	.17
LCOST(LOW.INC)	-3.90**	-6.14***	-5.47***
LCOST(MID.INC)	-3.92**	-6.20***	-5.48***
LCOST(HIGH.INC)	-3.94**	-5.87***	-5.02***
MOTH_USE	.77	1.45***	.94*
MOTH_HE	14.10	14.75	15.60
MALE	.04	1.92**	.60
HIGHER EDUCATION:			
Chi square of the model: 433.64***			
effect on	HNUE P_1/P_N	UNIV P_2/P_N	
INTERCEPT	11.95	8.24	
LOG(FULL INC.)	2.50	3.04***	
LOG(FAM.SIZE)	-1.75**	-1.59**	
LCOST(LOW.INC)	-7.54***	-7.35***	
LCOST(MID.INC)	-7.24***	-7.45***	
LCOST(HIGH.INC)	-7.16***	-7.07***	
MOTH_USE	- .04	.80**	
MOTH_HE	2.20***	1.53**	
MALE	.63	1.30***	

*: level of significance = 90%; **: 95%; ***: 99%

The three last conclusions together confirm the general impression that higher non-university education is more equally accessible than the university. More generally, our empirical results suggest that substantial financial tresholds still hinder the social mobility of youngsters from lower socio-economic backgrounds. At the same time, they point to the (potential) stimulating effect of various kinds of financial transfers like child allowances, tax deductions for families with studying children, and study grants.

1274

ALTERNATIVE ROUTES FROM VOVATIONAL EDUCATION TO THE LABOUR MARKET. LABOUR MARKET EFFECTS OF FULL–TIME VS. DUALIZED VOCATIONAL EDUCATION

Rolf Van der Velden, & Bob Lodder, Research Centre for Education and the Labour Market, University of Limburg, Maastricht, The Netherlands

Introduction

This paper addresses the labour market effects of two rather different forms of initial vocational education, full–time education with practice periods that have no employment status versus dualized education with an emphasis on continuous on–the–job/in–service training and employee status for the apprentice. Although most developed countries know both forms of vocational education, these usually have clearly different positions in a qualitative (standardization of curricula and degrees) and a quantitative (cohort shares) sense. Germany, for example, has a strongly developed system of dualized education that has a dominant position over its full–time counterpart. In the USA, on the other hand, most young people enter the labour market from full–time forms of education where vocational tracks have hardly developed.

The Netherlands present an interesting case, because in this country both systems coexist and contribute largely to the vocational schooling of each new generation. That gives us a good opportunity for investigating differences in the labour market effects of these 'treatments' in vocational education. In this article we address two questions:

- is there a difference in labour market position between graduates from fulltime vocational education and from the dual system, both shortly after leaving the educational system as well as after a period of 6 or 7 years?
- is there a difference in labour market careers (i.e. the chances of getting promoted or getting unemployed) between these two groups?

Theory

Our research uses a 'labour queue' framework (Thurow, 1975). Employers are supposed to rank applicants for a job according to their assessed training costs for the job. Not only individuals hold a position on the labour market, types of training can also be said to hold a position on the labour market. Glebbeek (1988) indicates that for large parts of the labour market – those sectors facing efficiency pressures – the relative position of a type of training is determined by the expected

1275

training costs of the graduates. With regard to the relative position of the full–time and the dualized system on the labour market, two training cost components seem particularly important: the extent to which graduates master the required qualifications, and their expected general learning ability. It can be stated that graduates of the dualized system generally have a better mastery of specific occupational skills, while graduates of full–time tracks have greater learning abilities.
Both components play an important role in the personnel selection process. While to some extent they can compensate for each other, the relative weights given to them by employers will differ between segments of the labour market. In so–called 'fchliche Teilarbeitsmrkte' (Lutz & Sengenberger, 1975) the emphasis in the selection process is on the mastery of specific occupational skills; in this segment graduates of the dualized system have an advantage, and (involuntary) mobility has no negative implications. In the 'betriebliche Teilarbeitsmrkte', however, the emphasis is on the long–term trainability of the labour force. In this labour market segment mobility is likely to have negative consequences for the occupational career; here graduates of full–time vocational education have an advantage. The same will hold for those segments of the labour market where rapid technological change takes place and therefore skills are subject to obsolescence.
Job search theories (McKenna, 1985) point out that the labour market entry of graduates is a turbulent period: it takes some time before a good match has been established. In this respect a major difference occurs between the full–time and the dualized system. For graduates of the latter, the period of job search precedes the start of their training course. In fact a good matchis a prerequisite for entering the dualized system. Therefore the occupational career in the first years after graduation will be more stable for graduates of the dualized system, and they are more likely to experience a good job match. In the long run these differences will diminish.

Research design
The occupational careers of some 1600 graduates from both full–time and dualized education were retrospectively investigated. All of them graduated in 1981 or 1982. A selection was made of equivalent vocational courses aiming at the following economic sectors: metal industry, building & construction, printing industry and health care. These branches represent different segments of the labour market and also show different rates of technological change.
Questionnaires were mailed to these graduates in the fall of 1987. The response was 69 %.

Model of analysis
We use event–history analysis (Allison, 1984) to examine the labour market careers of the different groups of graduates. Each individual career is broken down into different job spells. The following three types of transition between jobs are examined: rise of the level of the function (promotion), decline of the level of function (demotion), job change within the same level of the function (horizontal mobility). Apart from these transition possiblities, we examined the chances of getting unemployed. We used the Cox Proportional Hazard model to estimate the transition rates.

Findings

The main findings can be summarized as follows:
The initial occupational careers of full–time education graduates are less stable, and less often lead to a good job match. In the long run however these differences tend to diminish, although graduates from the dualized system keep a small advantage. Despite the fact that graduates of the dual system in general hold a strong position on the labour market, indications are found that they are more vulnerable to skills obsolescence. Beyond a certain age they run a greater risk of getting unemployed, while the chances of getting promoted diminish. Working outside the sector for which one is trained involves the risk of under–utilization, especially for graduates of the dual system.

References

Allison, P.D. (1984). Event history analysis. Beverly Hills: Sage.

Glebbeek, A.C. (1988). De arbeidsmarktpositie van opleidingen. Tijdschrift voor Arbeidsmarktvraagstukken, 4, p.75–89.

Lutz, B. en W. Sengenberger (1974). Arbeitsmarktstrukturen und öffentliche Arbeitsmarktpolitik. Göttingen: Otto Schwartz & Co.

McKenna, C.J. (1985). Uncertainty and the Labour Market. Recent Developments in Job Search Theory. Brighton: Wheatsheat.

Thurow, L.C. (1975). Generating Inequality. New York: Basic Books.

Velden, R. van der, D. Akkermans, H. van der Heul en F. van der Zee (1989). De lange arm van het onderwijs. Een vergelijkend onderzoek naar de lange–termijn effecten van opleidingen bij MBO en Leerlingwezen. Eindrapport. Groningen: RION.

VOCATIONAL SCHOOLS AND THE STRUCTURING OF INEQUALITY: A MALTESE CASE STUDY

Ronald G. Sultana, University of Malta, Malta

Malta's status as a British colony between 1800 and 1964 resulted in the stunting of its economic development (Clare, 1988) so that when the Island gained political independence, it did not have a sufficiently developed or diversified economic sector. A number of reports from the mid-twentieth century onwards had pointed out that unless Malta invested in technical education, then it would be very unlikely that industrial investment, by local or foreign capitalists, would take place. This advice could not be heeded because, as is common in countries with a colonial experience, the indigenous population aspired not for technical/manual occupations, but rather for positions with the civil service. The latter sector was considered to be prestigious mainly because of its association with the colonial rulers.

Technical education was introduced in the late sixties in Malta, but the secondary schools developed had an academic focus with a technical bias. It was only in 1972 that vocational schooling was set up by a Labour government, as part of a five-year economic development plan which had as a goal the facilitation of the process of industrialization. Perusal of the documents held at the Department of Education archives (Sultana, 1992) suggests that vocational schools at the secondary level were set up in Malta with at least three aims in mind. The first aim was economic: trade schools would train students to fill the new positions to be generated in industrial set-ups, at the operative and craft levels. The previous technical education provisions were denounced as being "technical in name only", and students in the new schools were to spend 75% of the time in workshops, and the rest in classrooms doing traditional academic subjects.

The second aim of the new vocational schools was to buttress the ideology of the Labour movement by giving status to manual and technical work. Inspired by the movements in China and Tanzania, the local Labour government set out to shake Malta's traditional hierarchial structures of power and privilege, arguing that the country's prosperity would not depend on the professions as much as on the labour of the working classes (Zammit, 1984).

A third aim of the new schools was to provide an alternative form of schooling for students who had lost motivation, or did not have an aptitude for academic schooling. A practical pedagogy and an education linked directly with the "real" world of work would revive lost interest in educational investment, and this would be to the benefit not only of students, but also to the country since potential would be developed rather than lost.

1278

Finally, a fourth aim was present in the setting up of vocational schools in Malta. This goal was implicit, and hardly ever treated at length in the official planning documents of the Department. Malta was feeling the effects of the world recession in the seventies. In addition to that, the closure of the British military base was imminent, and unemployment, which was already high, was expected to spiral even further upwards. In this situation, Malta chose to forestall the shock by keeping as many young people in schools as possible, for as long as possible, by raising the school-leaving age from fourteen to sixteen. The new trade schools, presented as being half-way between the traditional schools and the world of work, were to attract those reluctant students who would otherwise have rebelled on finding themselves forced to spend two extra years at school. In addition to this, the Government warned that nobody in Malta would be employed in skilled labour unless he or she had obtained a school-leaving certificate from the new schools. Data emerging from a four-year long research project on Trade Schools is showing that the economic imperatives which the new initiative had to respond to overshadowed the educational and ideological ideals expressed overtly at the planning stages of the whole reform. Pedagogical innovation has been minimal, and there certainly is no indication that "the hand" and "the mind" have been linked in the course of the transmission of knowledge. Rather, responses to a questionnaire given out to 680 students in their third year in these vocational schools show that:

- students generally choose to leave academic schools, rather than choose to attend trade schools. They end up in the latter because there is no other alternative to traditional schooling;
- by the end of the second year in these trade schools, absenteeism rates are as high as 32% for both boys and girls;
- close to 90% of all students in vocational schools are from a manual and working class background (Sultana, 1991).

In addition to this, tracer studies carried out by the Department of Education, by the author, or by students under the author's supervision (Sultana, 1992) have shown that large numbers of graduates of vocational schools do not use the skills they have learnt at school, because they find work in sectors of the economy for which there is no preparation. Thus, for instance, 30% of the male graduates found work in catering establishments in 1991, but there are no catering courses in vocational schools at the secondary level in Malta.

It would seem therefore, that even though precedence was given to the economic utility of the new schools, there are strong doubts as to whether that function is in fact being served. Local industrialists claim that the training offered in the schools is not responding to their needs, that the equipment and tools used there are out-dated, and that there is little synchronisation between developments in the economy and in the school curriculum. One key illustration of this is the bias of female trade schools in favour of industrial sewing, which, if we exclude from consideration the gender stereotyping that that entails, was a reasonable curricular response to Malta's investment in the textile industry in the early seventies. That industrial sector is now in decline, but the curriculum has not been reformed to take that change into account (Darmanin, 1992).

Given that there are clear signs that the economic function of vocational schools is suspect, the question arises whether these schools are fulfilling other functions

1279

which might justify their continued existence at a time when diversified schooling at that level is on the decline the world over. Indeed, comparative research on this topic has shown that:

- Industry is the best place for vocational training to be carried out in. Industry can train workers in vocational skills more quickly, more efficiently, and at less cost than schools. Small countries with few resources find it difficult to finance the cost of offering the newest equipment to students. In addition to that, the best engineers and craftsmen are more likely to be employed in industry than as teachers, because the wages are higher. Students therefore are thus at a double disadvantage, since the equipment they use is often out-dated, and the teachers they have not in tune with the latest technological developments.
- Diversified schooling is untenable not only because it fails to fulfil efficiently the economic aim it sets out to directly address, but also because it efficiently defeats a social aim which it overtly sets out to achieve. Vocational schools, and vocational streams in "comprehensive" type schools, have tended to attract students from the working classes, and to entrap them in occupational futures similar to those of the students' parents. In addition to this, vocational schools and streams tend to have lower status, and are utilised as a convenient dumping space by teachers in order to maintain the smooth functioning of the mainstream.

This is clearly the case in Malta, and the only function that vocational schools can be said to be fulfilling efficiently is the continued structuring and reproduction of class and gender inequalities through the different curricular experiences that they are offering. The charge that schooling serves as a sorting and channelling device is nowhere more true in Malta than for vocational schooling.

References

Clare, A.G. (1988). "Features of an Island economy", in V. Mallia Milanes (ed.) The British Colonial Experience: 1800-1964 , Mireva: Malta.

Darmanin, M. (1992). "The labour market of schooling", Gender and Education, in press.

Sultana, R.G. (1991). "Social class and educational achievement in Malta", in R.G. Sultana (ed.) Themes in Education: A Maltese Reader Mireva: Malta.

Sultana, R.G. (1992). Education as Human Capital: A Political Economy of Vocational Schools in Malta Mireva: Malta, forthcoming

Zammit, E.L. (1984). A Colonial Inheritance Malta University Press, Msida, Malta.

1260

SOCIAL CLASS, GENDER AND VOCATIONAL TRAINING: QUALITATIVE STUDIES OF THE EXPERIENCE OF YOUNG WOMEN

I. Bates, University of Sheffield, United Kingdom

This paper will be based on ethnographic research which took place in the context of the U.K. Economic and Social Research Council 16-19 Initiative * The purpose of these ethnographic studies was to examine the role of vocational training in relation to social reproduction. This led to a research design which included studies of a variety of groups on different tracks within education and training, allowing scope for consideration of both social class and gender as factors affecting occupational socialization, In adopting this approach e hare been influenced by the various additions in education research which illuminate inequality and stratification, and have drawn particularly on concept and insights from the sociology of education.

In order: to examine social processes within vocational training the researcher undertook participant observation studies with groups of young people on specific vocational training schemes over about two years. These studies revealed signigicant adjustments in career aspirations and orientation towards work in the course of vocational training. In this paper, which ;ll concentrate on the experience of young women, the central questions examined were:

- How did the young women involved come to enter their career?
- How did social class, fart i background and gender influence processes ofoccupational socialization?
- How did the above social influences interact with vocational training?

The analysis will focus on ways in which social class, family background and gender shaped trainees' perspectives and experience of training. It will contrast the official ideology of vocational training, which emphasizes equal opportunities and possibilities for individual career development, with the training processes observed. In practice these processes involved screening for cultural attributes, thus reinforcing class and gender-related divisions of labour. It will be argued that the role of training needs to be examined in relation to the wider soial skucures in which it is enmeghed and particulrly in relation to social class, gender and the labour market.

*The ESRC 1-19 Initiative was a programme of research, conducted y tens in the Universities of Sheffield, Surrey, Liverpool, Dundee and Edinburgh. It focussed on the development of "careers" and "identities". The programme was based on a longitudinal study of 5,000 young people and a group of ethnographic studies.

1281

References

Bates, I. (1989). Versions of vocationalism: An analysis of some social and political influences on curriculum policy and practice. British Journal of Sociology of Education, 10, 2, pp. 215-231.

Bates, I. (1990)."No bleeding whining minnies". The role of YTS in class and gender reproduction. Journal:of Education and Work, 31 , pp. 91-109.

Banks, M., Bates, I., Breakwell, G., Bynner, J., Emler, N., Jamieson, L., &, Roberts, K. (1991). Careers and Identities, 224 pages. Milton Keynes: Open University Press.

Bates, I. (1990). "No roughs and no really brainy ones": An wthnographic atudy of a BTEC National Diploma Course in Fashion Design. British Journal of Education andWork, 4, 1, pp. 79-90.

Bates, I. (1991). Closely obseved thing: An exploration of links between social structures, training and identity. International Studies in Sociology of Education, 1, pp. 225-243

Bates, I., &, IRiseborough, G. Thatcher's children: Youth in a Ccassless society? Milton Keynes: Open University Press, forthcoming.

1282

STUDENTS' DIFFERENTIAL TEXT PRODUCTION IN THE REGULATIVE CONTEXT OF THE CLASSROOM

Ana M. Morais, & Helena Antunes, Department of Education, Faculty of Sciences University of Lisbon, Portugal

Summary

The analysis of the classroom interaction has been fundamentally focussed on the instructional context undervaluing the crucial relations at the level of the regulative context. Teachers and students are not aware how the regulative discourse and practice have real existence in the school and embed the instructional discourse and practice marking ideologically the pedagogic practice. This has a number of consequences for children's achievement, particularly the disadvantaged.

The research reported in this article is focussed on the regulative context of the classroom and intends to analyse the differential text production by students with different backgrounds (social class, race) and submitted to three different pedagogic practices. These practices, differing in power and control relations, required from the students specific socio-affective dispositions. It was intended to study the extent to which children possessed the specific coding orientation at the level of a given regulative pedagogic practice and also to understand the underlying reasons. Bernstein's conceptual instruments of pedagogic code and also of recognition and realization rules were used to develop the analysis.

The study is included in a broader research carried out by the Project ESSA (Sociological Studies of the Classroom) with a sample of 80 students (10^{-} - 12^{+}) in a preparatory school with a fundamentally working class population of the suburban area of Lisbon. The students were divided in three sets each of which was submitted to pedagogic practices differing in terms of power and control, from a practice with strong classifications and framings to a practice characterized by weaker classifications and framings, that is from a practice centered on the transmitter to a practice more centered on the acquirer. The classes were socially similar and heterogeneous in terms of social class, race and gender. They were taught by the same teacher (female) who changed her pedagogic practice according to the school class. The main aim of this research is to find out pedagogic practice(s) which improve all children's achievement.

In the context of this broader research, an heterogeneous (social class, race) sub-sample of 6 children (2 from each pedagogic practice) was selected. They all had shown good/reasonable achievement in the instructional context. In the regulative context they represented discrepant cases, i.e. they were children who departed from the general pattern for their class/race: three white middle class children had

1283

shown poor achievement in the socio-affective dispositions required by the pedagogic practice to which they were submitted and three lower working class children had shown good/reasonable achievement.

A qualitative methodology was used including classroom observation, interviews and questionnaires. The observation of the classroom, made according to an etnographic perspective, gave direct data about children's regulative textual production (realization rules). These data were also used as the fundamental basis for the selection of the 6 children. The interviews, which were semi-structured, were devised to provide data about the possession of recognition rules; they also provided some data about the possession of realization rules. The data from the questionnaires gave information about children's affective dispositions towards the regulative text selected. The relation of all data obtained made possible to infer the degree of acquisition of recognition and realization rules in a specific pedagogic practice. Furthermore, it was possible to understand reasons which underly children's differential textual production.

The results suggest that the specific coding orientation in the regulative context is a function of the interaction between family factors and the pedagogic practice children receive. Disadvantaged children who were good achievers in the regulative context had in fact acquired the pedagogic code of a specific pedagogic practice, that is they had recognition and realization rules for that practice. Their motivations and aspirations together with self-compensatory strategies made them to produce the legitimate text even when it represented attitudes and values they did not share. On the other hand, the results also show that white middle class children who were bad achievers in the regulative context, had strong misconceptions about the regulative text required by a specific pedagogic practice. This prevented them of acquiring the recognition and realization rules to produce the legitimate text.

We believe that the approach followed in this research is original in its conceptual design and may constitute a contribution for a sociological study of attitudes, beliefs and values and their influence on students' achievement. Within the limits of the study, it was possible to appreciate the extent to which the production of the text required by the school in the regulative context is dependent on the possession of recognition and realization rules. It was also possible to unveil some underlying reasons for differential acquisition of those rules. From the study some directions for teacher's action can be derived.

References

Bernstein, B. (1977). Class, codes and control: Towards a theory of educational transmissions (2nd ed.). London: Routledge & Kegan Paul.

Bernstein, B. (1990). Class, codes and control: The structuring of pedagogic discourse. London: Routledge.

Morais, A. M. (formerly Domingos), Fontinhas, F. and Neves, I. (1992). Recognition and realization rules in acquiring school science: The contribution of pedagogy and social background of students. British Journal of Sociology of Education (in press).

1284

A DESCRIPTIVE STUDY ABOUT COMPULSORY SCHOOL DROPOUTS

M. Groppo, O. Liverta Sempio, E. Confalonieri, & G. Scaratti, Dipartimento di Psicologia, Università Cattolica del Sacro Cuore-Milano, Itali

Introduction

Studies into the dropout phenomenon consider both quantity (Cairns R.B., Cairns B.D., Neckerman H.J., 1989; Ribolzi L., 1990) and quality aspects. With regard to the quality ones, recent studies on school experience examine both individual dropout characteristics (Cairns R.B., Cairns B.D., Neckerman H.J., 1989) and relations between school structural aspects and dropout scholastic career (Malizia G., Frisanco R., 1981; Dasgupta B., 1987; Barrington B., Hendricks B., 1989; Bos K.Tj., Ruijters A.M., Visscher A.J., 1990).

The dropout phenomenon is considered as a developmental process that may begin in the earliest grades (Barrington B., Hendricks B., 1989; Cairns R.B., Cairns B.D., Neckerman H.J., 1989). Concerning this process J.D. Finn (1989) suggests two models of development: the "frustration-self-esteem" and the "partecipation-identification" models.

For a deeper understanding of the dropout phenomenon, it would be necessary to link it with the more general process of personal identity construction. The dropout experience in school occurs in fact during the adolescent period whose evolution steps consist of the construction process of personal identity (Erikson E.H., 1963; Blos P., 1979).

The present study deals especially with the early compulsory school dropouts. Its goal is to describe the "dropout condition" on the ground of perceptions and representations that the dropouts have on their scholastic, familiar, friendship and working experience.

Method

The examined population is that of the public compulsory schools of an industrial town nearby Milan. Within this population 76 dropout pupils (age:14 to 18) were found and examined both through quantity (grids and questionnaire) and quality (individual interview) instruments.

The data resulting from the questionnaire were worked out in terms of percentage distribution. As for the interviews, a thematic analysis of the content was made (Ghiglione R., Beauvois J.L., Chabrol C.L., Trognon A., 1980).

1285

Results and discussion

As for structural results, this study agrees with the relevant literature about dropouts social-family situation and their scholastic career (Malizia G., Frisanco R., 1981; Barrington B.L., Hendricks B., 1989; Censis, 1990; Bos K.Tj., Ruijters A.M., Visscher A.J., 1989).

Dropout life experience is marked by personal growth problems, typical of this age, which school makes more difficult.

Moreover, relationship with parents, especially with the father, seems difficult (clashing, overriding, emotionally detached): therefore, children feel demeaned, inadequate, not "identified" by their parents.

Not all the subjects seem equally troubled on an emotional, growing-up level because of these problems (Offer D., Ostrov E., Howard K.I., 1981; De Vito E., Palazzi S., Luzzati D., Guerrini A., 1989). Some of them avoid the conflictual relationship with the parents and assume a conventional, obliging attitude. Such attitude leads to the renounce of trying oneself and to the mistrust in the surrounding environment to help in the growing process.

Even towards future plans their attitude is not risking the present for an uncertain future. The subjects claim that nothing will change: same job, same friends, same freetime activities. This continuity makes them confident, but shows a kind of aquiescence, as if past and present cannot enable them to think of a real future (Winnicott D.W., 1974).

As for school experience, dropouts show a certain detachment from a period which is considered as closed on one hand, and which mostly evokes tiring situations. School does not seem to have helped them, either from a cognitive and relational point of view. Whereas in relation to school dropouts feel themselves negatively, in relation to work they express more positive feelings. Dropouts school experience therefore gathers frustration, self-esteem feelings and school identification process (see also "partecipation-identification model", Finn J.D., 1989).

The decision of leaving school was made by the subjects themselves: such lonely decision shows, once again, as the environment is unable to support their growth process. School, family and friends could not possibly help them.

From this point of view, adults negation of the subjects "immaturity" leads them to an early adulthood (Winnicott D.W.,1968).

As for the two education institutions, family and school, it arises the necessity of their greater presence in the dropout life, especially with a view to intervention.

References

Barrington B.L., Hendricks B. (1989). Differentiating characteristics of high school graduate, dropouts, and nongraduates. Journal of Educational Research, 82, 309-319.

Blos P. (1979). The adolescent passage.Developmental issues. New York: International Universities Press.

Bos K.Tj., Ruijters A.M., Visscher A.J. (1990). Truancy, drop-out,class repeating and their relation with school characteristics. Educational Research, 32,3, 175-185.

Cairns R.B., Cairns B.D., Neckerman H.J. (1989). Early school dropout: configuration and determinants. Child Development, 60,6,1437-1452.

Censis (1990). XXIV rapporto/1990 sulla situazione sociale del paese. Milano:F.Angeli.
Dasgupta B. (1987). Pupils' failure experience and dropping out. Psychological Research Journal, 11,1,17-23.
De Vito V., Palazzi S., Luzzati D., Guerrini A. (1989). Il Sè e l'immagine di sè nell'adolescenza. Età Evolutiva, 32,69-78.
Erikson E.H. (1963). Childhood and society. New York: W. W. Norton.
Finn J.D. (1989). Withdrawing from school. Review of Educational Research, 59,2,117-142.
Ghiglione R., Beauvois J.L., Chabrol CL., Trognon A. (1980). Manuel d'analyse de contenu. Paris: A. Colin.
Malizia G., Frisanco R. (1981). I fattori dell' abbandono nella scuola media. Risultati di una ricerca. Orientamenti pedagogici, 28,955-989.
Offer D., Ostrov E., Howard K.I. (1981). The adolescent: a psychological self-portrait. New York:Basic Books.
Ribolzi L. (1990). Le disuguaglianze di fronte all'istruzione:la base informativa. Scuola democratica, 1,73-80.
Winnicott D.W. (1968). Concetti contemporanei sullo sviluppo dell'adolescente e loro implicazioni per l'educazione superiore. In Winnicott D.W. (1974). Gioco e realtà. Roma:Armando,229-248.
Winnicott D.W. (1974). Gioco e realtà. Roma:Armando.

STUDENTS' GENDER AND SOCIAL SELECTION IN COMMERCIAL SECONDARY–LEVEL EDUCATION IN FINLAND

Marja–Leena Stenström, Institute for Educational Research, University of Jyväskylä, Jyväskylä, Finland

Introduction

During the last three decades the ideology of educational policy has been that everyone should have equal possibilities to succeed, regardless of his descent, gender, place of residence or other such factors. The goal of educational equality also has been to increase social mobility. The concept of equal educational possibilities differs in different conditions. We can say that there is no unambiguous meaning for equality.

The starting point of this study was the assumption that social selection is one function of the educational system. The study examined the realization of educational equality in vocational education on the basic line of commerce and administration covering both the institute–level and the school–level.

Problem

At the end of the 1970s a reform of secondary–level vocational education was initiated with the aim to integrate it into fewer basic lines. Students who start vocational studies after the comprehensive school are placed into one of the 25 basic lines for an introductory period of one year. After that they enter a line of specialization either at institute–level (higher) or school level (lower). The purpose of the reform was among other things, that students entering vocational education would choose their *field* of education before choosing the *level*. In Finnish educational discourse the concept of such a *two–stage selection* system refers to the system in which the basic line is selected on application to vocational education (1st stage) and the specialization period at the end of the one–year introductory period (2nd stage).

In this paper the study sought answers to the following questions:

- Does there exist a relationship between the selection of the school system and the social background?
- Does there exist a relationship between the selection of the school system and gender?

1288

Method

The population of the research reported here included all the Finnish institutes of commerce and administration. The population was divided into four strata: three regional areas, and one stratum consisted of institutes which had national spezialized lines (some specific lines, for example, data or library lines, are located at only a few institutes). The sampling method was PPS–sampling (probability proportional to size). The sample of this research includes fourteen institutes and about 1800 students. The main research data were collected by means of questionnaires in 1984 – 1986 in the different stages of the follow–up research.

Result

One third of the students were men and two thirds were women. The achievement level of the starting students had been average in the comprehensive school. Over half of the students had transferred straight from the comprehensive school, twelve per cent had finished comprehensive school a year before and over one third earlier than that. More than a third had worked before they began to study, and similarly one third had also other futher education. Most of the students (91 %) on the line of commerce and administration had been able to realize their goal, with this particular line having been their first choice. Only six per cent had been seeking admission primarily to nursing institutes and three per cent to other institutes.

Log–linear models were used to analyse one part of the collected data (Kennedy 1988). One way to describe the adequacy of the model is to count the odds ratios from the expected frequencies in log–linear model (Agresti 1984).

The relationship between the social background and the student's goal level is maintained when the school system has made students selection. Of course the goal level affects how the desired level will come true. Students suggest desires of the specialized lines at the end of the introductory period and then the establisment makes the students selection on the basis of the achievement of the introductory period. Regardless of how the students succeed in the introductory period, the children, whose fathers' basic education is the lower secondary school, get addmission to the institute level one and a half times (1.5) more often than those whose fathers' basic education is the elementary school. When the fathers' education is the upper secondary school, the children get admission to the institute level almost three times (2.7) more often than those whose fathers' education is elementary level. The two–stage selection system does not remove the connection between students' social background and the admission level, although the achievement in the introductory period as a criterion of the specialization period equalizes the situation.

The admission of the institute level according to gender varies in terms of the achievement in the introductory period. Men with a low achievement grade were selected by an establisment three times more often to the institute level than women. Correspondingly, men with a medium achievement grade were selected about almost two times more often than women and when the achievement is at a high grade men were selected almost seventeen times more often than women. Of course, such a great difference has arisen from the fact that the desired level between men and women has become differentiated. Men have applied almost

three times more often to the institute–level, regardless of their achievement in the introductory period. Furthermore the result is a concequence of small frequencies, nevertheless so the result indicates that there exist differences in admission to the institute level between men and women.

Discussion

Educational selectivity is, however, not a phenomenon that can be removed only by means of education. Several investigations have found that the school cannot compensate for what is caused by the other institutions of society, such as family, culture and workforce. The school can try to reduce differences, which are due to social background and gender, by means of the curriculum content which cannot be ignored, when the school system is reformed.

When, in general, the parents try to ensure for their children almost their own level of education, it is likely that the effect of the parents' educational level is repeated in the choices of their children. When the selection occurs in accordance with social background, it also means, on the other hand, that education has significance in terms of social mobility over the generations.

References

Agresti, A. (1984) Analysis of ordinal categorial data. (New York: Wiley).

Kennedy, J.J. (1988) Applying log–linear models in educational research. Australian Journal of Education, 32, (1), pp. 3–24.

A DYNAMIC ANALYSIS OF THE TRANSITION FROM EDUCATION TO EMPLOYMENT: THE CASE OF LUXEMBOURG

Laure-Suzanne Allegrezza-Carvoyeur, Lasare Université de Nancy, Nancy Cedex, France, & Jean-Luc Kop, CEPS/INSTEAD, Walferdange, Grand-Duché de Luxembourg

Abstract

The objective of this research is to analyze the transition between school and employment in a dynamic perspective. The duration of non-employment after leaving school is the dependant variable. Survival analysis models which allow uncertain knowledge on non-employment duration (some non-employed persons have not yet a job at the end of the registered period) are performed. The data are taken from the four first waves of the "Luxembourg Panel (PSELL) of Income Dynamics", and concern 201 subjects out of the 6000 subjects in the whole sample. Individual situations are recorded on a monthly basis, which allows us to compute a precise non-employment duration. Finally, different kinds of regressors (individual and household variables) are introduced in order to find the best determinants of the transition.

Youth unemployment in Europe is one of the major problems governments have to manage. Even in a country such as Luxembourg where unemployment rate is very low (less than 2 %), young people have more difficulties than others to find a job. More specifically, all young people do not immediately find a job after leaving school. Several *a priori* explanations can be advanced: some of them do not search for a job, others may have no qualification or have qualifications which do not fit the labor market. Theses explanations are certainly insufficient: so, we have to find which factors influence the transition between education and employment. For this kind of study, longitudinal data are particularly appropriate. They can help solving questions like:

- which are the flows for the school leavers (e.g. evolution of the level of qualification of school leavers, evolution of duration between end of school and beginning of first job, evolution of salaries...)?
- What kind of indicators can account for such evolutions and how can tendencies be modelized?
- How can the educational system be evaluated in terms of adequation to the labor market?

1291

- Which individual and household variables influence quality of transition between education and employment?

In this study, we will only focus on this last objective.

For this purpose, we use micro-economic and longitudinal data from the four first waves (1985 to 1988) of the "Luxembourg Panel (PSELL) of Income Dynamics". In this database relevant information is recorded monthly so that it allows a dynamic perspective.

Of the 6000 individuals in the survey, 217 subjects left school during the observation period. 13 have been removed of the sample because their situation was too specific (prsons with mental or physical handicap, in military school or who go back to school after a long interruption).

In this study, the quality of the transition between school and employment is operationalized by the *duration* between the end of school and the beginning of the first job rather than by the probability of having or not having a job. This choice is justified by the following arguments:

- the analysis of the probability of having a job can be managed with traditional regression methods like logit analysis but this kind of analysis assumes a stability of situations (static point of view)
- analysis of duration of non-employment after leaving school can be managed with survival regression models. Such models can take into account censoring times (i.e. unknown durations due to some persons who did not find a job by the end of the observation period), and time-varying explanatory variables.
- survival regression models allow hypotheses on distribution's shape (Weibull, exponential, log-normal...)

Most results of traditional analyses of transition point out the influence of individual variables on the duration of unemployment. That's also an aspect of our hypotheses. But our hypotheses take also into account household variables because we think that family can have a great impact on professional insertion. Different kinds of explanatory variables are introduced in the regressions in order to find the best determinants of the transition between education and labor market. Individual variables included are : sex, age, kind of vocational training (e.g. technical *vs* general education), level of school attainment (e.g. number of school-years) and individual socio-cultural integration. This last variable is particularly interesting in the case of Luxembourg. Indeed, two typical aspects of this country are a very high rate of immigration, and the necessity to speak and understand several languages (especially French and German). The indicator used is based on several variables : nationality, abilities in Luxembourg, French and German languages.

The socio-cultural integration of a person may be very different from the integration of his household. So it is also important to take into account the household's integration. Another explanatory variable is social contacts which can have an important impact on the transition (many relations may facilitate opportunities to find a job). An index of social contacts (with family, neighbours, friends, colleagues and with professional or cultural associations) is introduced in the regressions. Socio-economical status of households and deprivation indicators are traditional variables used in studies on education. However, the social indicators research has pointed out the difference between objective and subjective indicators and the interest of the last ones. In this study we take into account these results by the use

of both objective socio-economical indicators and subjective socio-economical indicators. In the first set of variables, we use school attainment of head, different kind of resources, income *per capita*, housing comfort. In the second set of variables, we use an index of material subjective difficulties and an index of subjective housing comfort. Correlations between subjective and objective indicators are only moderate.

Some interesting results appear from preliminary multivariate analyses of only 78 individuals from two waves (1985-1986) . First, in spite of a very low rate of unemployment in Luxembourg, some people did not find a job during the observation period. Second, some individuals who left school before the age of compulsory school attendance (16 years) have great difficulties in finding a job. Third, and surprisingly, several classical variables used in similar studies are not significant (sex, nationality, and so on). Fourth, other criteria have unexpected effects: for example, people who leave the technical vocational system have more difficulties in entering the labor market than people who leave the general educational system.

These preliminary results will be completed by studying the transition between school and employment with a larger sample, using a longer period of observation and using other explanatory variables as listed above. We believe this approach can adequately predict inequalities of integration in the labor market, and can be used to minimize them.

YOUNG PEOPLE'S DEVELOPING LIFESTYLES AND THE TRANSITION FROM SCHOOL TO SOCIETY: A LONGITUDINAL STUDY

L.B. Hendry, J. Shucksmith, J.G. Love, & A. Glendinning, Department of Education, University of Aberdeen, Scotland, United Kingdom

Introduction

How do various adolescent lifestyles develop in the transition from school towards adulthood? Our understanding of the development of adolescent lifestyles needs to take account of a 'network' of determinants. These are common to all concepts of lifestyles and include the individual with his/her socialisation process and present personality as well as his/her immediate surroundings (family, teachers, friends, workmates, colleagues and so forth). It also includes connections with the social and economic environment. Thus objective conditions and subjectives perceptions are contained within one model. In particular, the linkages of this 'network' help to formulate a better understanding of the development of adolescent lifestyles in the transition from school to society.

Method

An ongoing longitudinal research project (Hendry *et al* 1985-92) is studying young people's leisure and lifestyle over an extended period. The project has been designed to operate at two complementary levels. There is a survey element and this has involved a sampling frame of 126 schools to represent a nationwide sample of Scottish young people. An initial survey of 10,000 young people from 10-20 years of age and comprising six cohorts was carried out in 1987. This was followed by two further large-scale questionnaire-based surveys of adolescents' socialisation and leisure in 1989 and 1991.

The second element of the project is a smaller, more intensive examination of a sub-sample of 250 young people living in eleven different socio-geographical areas of Scotland. This ecological approach is intended to examine some of the implications of social environment on young people's lifestyles. The overall purpose of the research project has been – and is – to develop an explanatory picture of young people growing up in modern society by the use of a variety of reasearch techniques and styles.

In this paper we focus on the first two sweeps of the large-scale longitudinal study and attention is restricted to one cohort of roughly 350 adolescents. The young

1294

people from this cohort were aged 15-16 years in 1987 and 17-18 years in 1989. This cohort is of particular interest because in 1987 these young people were all still attending secondary school (but at the end of compulsory school education) whilst by the time of the survey in 1989 they had all left school and had begun the transition to adult society.

The approach taken using factor and cluster analyses enables us to examine the developmental context of lifestyles during adolescence. In order to focus in some detail on aspects of the adolescent transition, young women aged 15-16 in 1987 are particularly considered. (These young women were still at school in 1987 and so we were able to follow them from compulsory secondary education to adult society and the world of work in 1989.)

Results

The results of these analyses are used to describe how a variety of personal, social, education and occupational influences interact in the development of lifestyles in the transition from school to adult society. In brief, the results show:

- inter-related and ecological aspects of living conditions produce a clear class-based differentiation of adolescent lifestyles;
- considerations of social class alone do not account for the clusters obtained. Differences in lifestyles also emerge within social class boundaries. We need to take account of 'within social class' differences in any real understanding of adolescent development;
- Importantly some lifestyle patterns are common to both middle and working class young women;
- When we look at the longitudinal transitions made by young women between the period when they are still at school and when they have left school and moved into adult society, identifiable lifestyles emerge. Yet the development of these clusters over time represents an exchange between clusters within social class boundaries and in terms of fragmentation across class boundaries the exchange is fairly minimal. The migration of individuals from one type of lifestyle to another is predominantly within social class boundaries;
- Young people in higher education, working youth and those young people on government training schemes or the non-employed do have different lifestyle patterns;
- When we look at young men in the 15-16 year old cohort in 1987 and again in 1989 we find some similarities with young women's lifestyles – but equally differences emerge.

Discussion and Conclusions

In discussing the main findings in relation to Colemen & Hendry's (1990) focal theory it is argued that the changing patterns of lifestyle development do not necessarily relate directly to focal theory. However it is possible to note a combination of life events, relational issues, values, attitudes and leisure interests which interact as focal theory suggests. This allows us to consider not only social class and gender differences but also important lifestyle variations within social class boundaries as young people make the transition towards adulthood. It is therefore possible using cluster analysis to gain clearer insights into the various aspects of attitudes, meanings and behaviours which go to make up more

conventional adolescent lifestyles within social class boundaries and to acknowledge the variations within social classes which produce more 'youth culture' oriented living patterns. Additionally, 'achievement-oriented' academic lifestyles do seem to comprise of different types of psychological stress, social relationships and life events between working class and middle class adolescents in the transition from school to adult society (Hendry *et al*, 1992).

References

Coleman, J C & Hendry, L B (1990). The Nature of Adolescence. London: Routledge.

Hendry, L B *et al* (1985-92). Young People's Leisure and Lifestyle project (Health Promotion Research Trust), Department of Education, University of Aberdeen.

Hendry, L B *et al* (1992). Young People's Leisure and Lifestyle. London: Routledge (forthcoming).

WORKING PRACTICES AND CHILDREN'S APPLICATION OF SUBJECT KNOWLEDGE IN THE PRIMARY SCHOOL

Stephen Bristow, & Charles Desforges, University of Exeter, United Kingdom

Background

Much emphasis is currently placed in classrooms on children learning through work which requires them to apply existing knowledge in order to learn from new experiences.

In British primary schools (where the research to be reported is being conducted) this approach to learning is implicit within both topic-work and more specific problem-solving activities. Within the National Curriculum guidelines (D.E.S., 1989) explicit mention is made of the part that knowledge application has to play in children's learning.

The key role that *old* knowledge plays in the acquisition of *new* knowledge is widely acknowledged among researchers in the field of cognitive psychology. However, the mechanisms of this role are not fully understood (Bereiter, 1985). In addition, attention within educational research is turning away from purely *in the head* explanations of learning to include the role that *culture* plays in cognitive phenomena (Säljö, 1991). The concept of culture too, however, remains unclear as does the relationship between it and cognition.

Advances in understanding these processes would contribute considerably to the design and management of classroom work.

The study reported here attempts to make such a contribution. It is part of a large project (n=48) on children's application of subject knowledge. The aims of the project are to identify how children apply knowledge (broadly conceived to include subject matter, general knowledge, knowledge of working practices and strategic knowledge) in learning from school work, and how children's strategies for applying knowledge develop in the age range 7 to 11 years.

The Study

The particular aim of this study was to explore how 7 year old children apply their knowledge of the directional nature of number in a novel and difficult task (specifically a problem involving movement between floors - above and below ground - in a block of flats).

Seven children aged 7 years comprised the sample, in two schools. Data were collected in two phases:

- Rich descriptions of classroom observations during mathematics lessons were made (focussing on child-child and child-teacher interactions, children's working practices - eg. their perseverance, use of specific apparatus, preferences for obtaining help and the nature of that help); examples of children's work were collected, and each child was interviewed to establish his/her existing knowledge, skills, goals and feelings relating to the content area of directional numbers and the working practices related to mathematics in the classroom.
- Each child then took part in a conversation with a researcher conducted around a verbal problem posed in relation to a hypothetical block of flats. Specifically each child was challenged to engage with the question "If you started on the second floor of a block of flats and went down six floors where would you end up?" Using techniques of critical exploration operated within a semi-structured interview, a protocol relating to the child's work on the problem was elicited. Each interview was audio-recorded.

The structure for the interviews was derived from our broadly conceived theoretical framework and included questions which probed for:

- the child's attempts to relate this experience to prior knowledge (both subject knowledge and general knowledge) and the selections made therein.
- the child's attempts at monitoring and evaluating his/her thought processes.
- the child's affective responses to the problem and the setting in which the problem was posed.
- the child's goals (both intellectual and social) in relation to the problem and the setting in which the problem was posed.
- the outcomes of the child's work on the problem in terms of newly acquired knowledge and/or knowledge restructuring.

Analysis

Using the technique of contrastive case study each data set was analysed in order to examine the child's application of domain and context-specific knowledge and the constraints placed thereon.

Findings

The findings will be presented in the form of an illustrative analysis contrasting two case studies. Issues relating to method, theory and practice will be discussed particularly as they apply to the effect of working practices on learning.

This research is supported by the Economic and Social Research Council (E.S.R.C.).

References

Bereiter, C. (1985). Toward a solution of the learning paradox. Review of Educational Research, 55, 2, 201-226.

Department of Education and Science (1989) Mathematics in the National Curriculum (Non-Statutory Guidance)

Säljö, R. (1991). Culture and Learning, Learning and Instruction, 1, 179-185.

INTERNATIONALIZATION OF EDUCATION

J. van der Pluijm, IVA, Tilburg, The Netherlands

Description of problem

In the field of primary and secondary education, the Dutch government wishes to encourage "...initiatives to implement bilingual education ... in so far as these do not involve additional expenditure and are freely accessible to all pupils" (O&W [Ministry of Education and Science], 1991: 14). This means that all schools will be able to provide education in a second language. Apart from the existing forms of bilingual education, namely Internationally Oriented Education (IGO) and the so-called European schools, a new form of bilingual education can be developed. This will be freely accessible, so that all pupils will in future be able to choose between education through the medium of Dutch and bilingual education.
The possibility of choosing bilingual education does not of course mean that every pupil will actually do so. From the point of view of policy-planning is it interesting to investigate what the nature of the future population of pupils in bilingual education may be. The present paper gives a description of a possible population. The problem for study can therefore be formulated as follows: which pupils will choose bilingual education and in what way will they differ from pupils who do not choose this new form of education?

Framework of description and questions studied

Dronkers (1991) distinguishes two perspectives by means of which it is possible to describe pupils' choice for bilingual education. The first perspective, the "cosmopolitan perspective", is taken from Ultee (1989). European unification means that the upper classes of society in the various countries are fusing together into a cosmopolitan culture. The wish to belong to that cosmopolitan culture is a reason to choose bilingual education. The second perspective, the "social class perspective" is taken from Boudon (1979) and Mare (1981). They distinguish decision points in education which are important for differences in educational opportunity. Children from the higher social classes have a better chance than those from the lower classes to make a more advantageous choice, that is to say to choose the highest level of education. Bilingual education introduces a new decision point, thus making it possible for the educational distance between the higher and the lower classes to be preserved. The choice of bilingual education can in other words be described from a "cosmopolitan perspective" and from a "socio-economic perspective". The problem to be studied can thus be formulated in terms of these two perspectives in the form of the following two questions:

- Do pupils who choose bilingual education differ from pupils who do not do so, in terms of the cosmopolitan and socio-economic backgrounds of the family?
- To what extent can the cosmopolitan and socio-economic backgrounds of pupils provide a basis for predicting the type of education which they will choose?

Data and method

Strictly speaking, it is not possible at the present moment to answer the questions posed for study. Bilingual education is, after all, still only a proposal. For many pupils it is still something which they feel is a long way off. It is however possible to give an indication of the answers, if one takes as one's starting point an existing form of bilingual education. Data has been collected on Dutch pupils who are taking the first year of the International Baccalaureate (IB) and on Dutch pupils at the same schools who are in the fifth year of "Pre-university secondary education" (VWO). Both types of school are well comparable as regards the starting and finishing levels (Van der Pluijm, 1990, 1991).

From the total group studied, a group of pupils was selected who have a real choice between regular and bilingual education:

- these pupils all meet the admission requirements of bilingual education to the same extent (that is to say, they have all lived or been to school abroad);
- the home language of these pupils is Dutch;
- these pupils attended a school in the Netherlands last year.

The pupils differ in the choice actually made. Some of them chose bilingual education, the others education through the medium of Dutch. The data collected relates to general features (e.g. sex and age), socio-economic background features (e.g. socio-economic status, father's level of education, mother's level of education, number of books owned), cosmopolitan background features (e.g. residence and employment abroad, visits abroad by parents in connection with work, foreign holidays), motives for choosing or not choosing bilingual education and the features with respect to the desired future after secondary education. Descriptive statistics were used in order to arrive at an answer to the first question to be studied. The second question was answered with the aid of logistical regression.

Results

Pupils in bilingual education differ from pupils in regular education in a number of features. The cosmopolitan background features show differences: pupils in bilingual education have visited more continents, their fathers go abroad more frequently and for longer periods in connection with their work, and their brother(s) and/or sister(s) have always had bilingual education also. There are also differences in the socio-economic background features: pupils in bilingual education come, on average, from families of higher socio-economic status. There is no difference between the two groups of pupils as regards age and sex. In short, pupils in bilingual education score higher than those in regular education on both cosmopolitan and socio-economic background features. The choice of bilingual education appears to be explained above all by the cosmopolitan background features of the pupil's family. The socio-economic background has only a minor

effect. It is striking that sex is also a significant factor: boys choose bilingual education more frequently than girls.

Conclusion and discussion

The choice for bilingual education appears to be influenced above all by the cosmopolitan background of the pupils. The socio-economic background also plays a role, but its influence is much less. These results must be viewed in relation to the composition of the group studied. All the pupils come from the higher social classes and all have lived or gone to school abroad. In view of the fact that, even so, the cosmopolitan factor still has a significant effect, it can be concluded that the cosmopolitan background is of great importance.

In conclusion, encouraging initiatives for bilingual education which is accessible to all pupils could lead to only those pupils being reached who come from families with both a high level of cosmopolitan orientation and a higher socio-economic status. This can accelerate the process of cosmopolitanization identified by Ultee among the upper social classes. Whether or not this process is desirable is a political question which goes beyond the scope of the present study.

References

Boudon, R. (1979). L'inégalité des chances. La mobilité sociale dans les sociétés industrielles. Parijs: Armand Colin.

Dronkers, J. (1991). Sluipende internationalisering. Sociale Wetenschappen, 34, 171-187.

Mare, R.D. (1981).'hange and Stability in Educational Stratification. American Sociological Review, 46, 72-87.

O&W (1991). Grenzen verleggen. Nota internationalisering van het onderwijs. 's-Gravenhage: Sdu.

Pluijm, J. van der (1990). Engelstalig onderwijs in Nederland, doctoraalscriptie, Tilburg: IVA.

Pluijm, J. van der (1991). Engelstalig onderwijs in Nederland, in: Dronkers, J., Meijnen, G.W. & H. Dekkers (red.), Onderwijs en Samenleving, SCO, Amsterdam, p.195-206.

Ultee, W.C. (1989). Economische eenwording, groter wordende sociale afstanden, blijvende culturele verscheidenheid? Afwijkende antwoorden op nieuwe vragen over ongelijkheid en cohesie in de Europese Gemeenschap sinds 1957 en na 1992, oratie, Nijmegen: ITS.

1301

BEST COPY AVAILABLE

LEARNING MOTIVATION, INTELLECTUAL AND EMOTIONAL DEVELOPMENT OF ADOLESCENTS BEARING DEPRIVATION

A.G. Moskovkina, & A.A. Sagdullaev, Institute of Defectology, Academy of Pedagogical Sciences, Moscow, Russia

Present day professionals in the field of education in Russia often face problems related to early deprivation although people standing somewhat apart from our situation may underestimate significance of this issue in education. The number of children of this deprivated category is growing yearly in contemporary Russia. Over 1 million of orphans and children deprived of parental care are being upbringing and educated in state institutions. Moreover milder forms of deprivation are discovered in choldren from low social classes with low level of income and education. Nowadays even children from middle classes suffer from deprivation to some extent. Deliquency and prostitution are going up rapidly. Social programming is necessary to promote education and social adaptation of children facing severe psychic deprivation, but subsidy is cut off. That is why the focus of our study is upon conseqtlences of deprivation in relation to health and peculiarities of emotional and intellectual development in two maior groups of heavily deprivated children; social and true orphans. The first group consisted of adolescents upbringing in state institutions from early age or from birth because their parents alcoholics with asocial behavior didn't take care of them. Symptoms of so called "alcohol fetal effect" were infreqluent in this sampe. That is retardation of phisical and psychic development, neurological and psychopathological disturbancies and specific pheotypical characters. The second group included adolescents from Afghanistan who lost their parents during military operations and are being educated in Tashkent, Uzbekistan. They have never been to any state institutions since lived earlier with their remote relatives. They were to return to their motherland after education in the former USSR. Control group of adolescents comprisinh those living with their parents was matched with deprivated groups in age and academic progress. L.S. Vygotski that psychic development of a child represents unity of biological and social factors served theoretical basis of our study. Then normal and abnormal development are submitted to the same rules. Age dependence of complex relationships between social and biological factors was taken into consideration as well. Adolescent age was chosen for study because of its critical significance for process of socialization in adulthood. In Russia orphans after 15 years old live independently without advice or consultation. Possibility of such independency is determined by their health, levelof intellectual development, emotional stability and professional skills

1302

Developing learning motivation is of significance as well. Medical and social inquiry, R.Kettel questionary, Milman's method of multisided studies of personality were used. Analysis of results indicated that these two groups of deprivated adolescents created rather different psychological and educational problems. Social osphans suffered from emotional disturbancies. They were astenic in their emotional preferences and frustrating behavior. They felt it difficult to get out with their troubles. On the contrary true orphans were stenic e ıough in their emotional preferences and frustrating behavior. The most strilking results were obtained from comparison of two groups of social orphans: mildly mentally retarded (IQ 70) and normal. The last had verbal intellect adequate to boderline mental re tardation. Apparently early emotional deprivation because of disturbances of mother-child relationships affected their speech development and verbal intellect as well. In spite their relatively poorer lknowlege of Russian afghanish orphans had higher verbal coefficient than their russian peers. Thus late deprivation of maternal care is less significant for emotional and intellectual development of a child if ever. Thus our data confirm that emotional deprivation at 6-12 month age (sensitive period) influence both emotional and speech development of a child. Lack of intimate personal contacts with adult at 1-2 years age ia of particilar significance for speech development. Our results indicated low learning motivation in adolescents-social orphans. There was difference between boys and girls. The latter had lower level of learning motivation than the first. Girls-orphans had low level of general every day motivation as well. There were common features in two groups of orphans possibly caused by similar conditions of education in state institutions of closed type . All of them had low motivation of communication. Orphans in closed institutions have very limited possibili ties of communication: teachers and peers not always to their liking. These contacts do not promote development of communication skills . Communication in closed institutions is not obligatory condition for satisfying one 's needs. All of the orphas independently of their intellectual level proved to be timi d and socially inactive . These features prevented them from positive social adaptation and integration in their adulthood life. Those features are considered as effects of education in closed institutions with rigid system of prohibitions and supression of any initiative. To social orphans attention deficit, hyperactivity, exitability and pathological forms of behabior were inherent . Most of them suffered from organic birth injury which was somewhat compensated later. Decompensation in puberty was infrequent. Some of them needed medical and psychological correction.The last was not always available. Overmedication and hospitalization groundless were infrequent. Identifying, subtyping, differential programming. Expanded intervention and prevention should be carried out. Moreover new ideas or conceptualizations that make these problems go away are necessary.

THE INFLUENCE OF PARENTS' EXPECTATIONS ON SCHOOL CAREERS

Anna A. van der Hoeven-van Doornum, Paul Jungbluth, Institute for Applied Social Sciences, & Marinus J.M. Voeten, Department of Educational Sciences, University of Nijmegen, Nijmegen, The Netherlands

It is generally known that schoolcareers are related to socio-economic status. SES has an effect on learning achievement in primary schools. And, apart from that, SES appears to play an additional role in the process of assigning children to types of secondary education. These effects of SES can partly be explained by factors related to cultural differences in the social background of pupils and by factors related to school characteristics. Recently it was shown that aspiration levels teachers set for their pupils also contribute to the explanation of these SES effects (Van der Hoeven-van Doornum, Voeten & Jungbluth, 1990; Van der Hoeven-van Doornum, 1990). Aspiration levels set by teachers for their pupils are mainly based on perceived learning performance and social background of the pupil. Aspiration levels are general goals the teacher expects to attain with the pupil. As a consequence of their expectations, teachers anticipate on attainment of goals hold for each pupil and by doing so teachers will, to a certain extent, achieve these goals in the end. It is to be noted that aspiration levels do not include specific learning contents of reading or mathematics the pupil has to master. A distinction between two kinds of aspiration levels must be made: there are absolute and relative goals. Absolute goals refer to what the teacher expects to be ultimately attainable for a pupil, such as the level of secondary education. Relative goals are related to instructional and educational processes, and will be indicated as attainable cognitive goals. Examples of cognitive goals are 'high demands for this pupil and emphasis on reading and mathematics'. The distinction between absolute and relative goals is of importance, because relative and absolute goals have different meanings and are not necessarily coupled. Parental expectations are probably one of the main characteristics of parenthood and education. In raising their children, parents express their expectations in many implicit and explicit ways, depending on their social and cultural background. Expectations of parents may work out in the same way as aspiration levels of teachers. In daily activities - like playing, reading, talking with their children - parents may also require high demands, thereby stimulating the cognitive development of children. Like teachers, parents may have absolute goals expressed as an explicit preference for a certain type of secondary education for their children. In this study

we focus on the effects of the latter type of parental aspiration levels of parents on school careers. The aim of this study is to investigate the influence of the expectations parents have for their children on learning achievement and by that on schoolcareers.

Research questions
In this study learning achievement at the end of primary school and the principal's advice about the type of secondary education for the pupil were chosen as the main variables to be explained. We were specially interested in the extent to which preferences held by parents for their child to enter a certain type of secondary school, are related to learning achievement, to aspiration levels set by teachers, and to advised type of secondary education. Parental preferences were elaborated as two variables which were analysed seperately. The first variable was whether or not parents had explicit preferences for their child's secondary education. And if they had, the second variable was defined to be the type of secondary education they actually preferred for their child. Aptitude and prior learning achievement were variables to be controlled for.

Model specification
Point of departure was a model adapted comes from a previous study in which models for school careers including the effects of SES, aptitude (IQ), learning achievement and teacher expectations on advised type for secondary education were analyzed (Van der Hoeven-van Doornum, 1990). From this initial model two other models were developed to studt the effects of parental aspirations, one for the existence op parental preferences and one for the aspired level of secondary education. It is our hypothesis that aspiration levels of parents have a mediating role in the effects of SES on learning achievement, teachers' aspiration levels the principal advice. As a consequence the model depicted in Fig. 1 has to be expanded to include an effect of SES on parental aspirations and effects from the latter on learning achievement, teachers' aspiration levels and the principals advice. After finding a well fitted model, we can asses to what extent effects of SES on school careers can be explained by parental aspiration levels.

Data collection
The study was carried out with 11- and 12 year old pupils in grade 8 of 53 schools for primary education in the eastern and southern part of The Netherlands and their parents. An extensive questionnaire on aspiration levels was filled out by the teacher for each individual pupil. In addition, there were achievement tests and an intelligence test for the pupils. At an early stage (in grade 7) information on the socio-economic status and parental aspiration levels was obtained from the parents. Further information on variables will be presented in the full paper.

Results
To answer the research question we applied the Lisrel procedure to covariance martrices. As expected, the initial model (1) without the preference of parents fit chi^2=8.39, df=5, p=.14 and AGF=.98 (N=701). The variances of the test score and the advice at the end of grade 8 were quite well explained by the model (77% and 86% respectively). About 70% of the total effects of SES on test scores and

advicecould be explained through learning achievement and teachers' aspiration levels at the beginning of grade 8. The second model was developed to analyse parental aspirations by adding the variable 'existence of explicit aspirations (1=preference, 0=none) to the initial model. With a slight modification, model 2 fitted also adequately to the data: chi^2=6.88, df=6, p=.14 and AGF=.98 (N=701). The existence of parental preference could not be explained very well by SES and IQ (7%). Comparing the result of the second model with the initial model, only a slight increase of the indirect effects of SES was noted (see Table 1, model 2). Twenty-two percent of the parents (N=157) could not indicate their preference for a certain type of secondary education for their child. As a consequence 544 parents were left for the third model. Children whose parents indicated an explicit peference were significantly higher on both SES and IQ. The third model was developed by adding the variable 'choice' (range 10-40, mean 26.40), that is the specific type of secondary education parents prefer for their child to model 1. Again with a slight modification, model 3 fitted the data well: chi_2=6.88, df=7, p=.44, AGF=.98 (N=544). The variable indicating the choice of parents is explained by SES and IQ for 35%. By adding the choice of parents the direct effect of SES on the test score at the end of the year disappeared completely and a substantial decrease occurred with respect to the direct effect of SES on advice (see Table 1, model 3 and Fig. 2).

Table 1

Direct and indirect effects of SES on learning achievement (test score) and recommendation at the end of grade 8 (as percentages of total effects) and observed correlations.

	model 1 (initial)		model 2 (preference)		model 3 (choice of parents)		r
effects	direct	indirect	direct	indirect	direct	indirect	
relative goals	.00	.00	.00	.00	.00	.00	.24
absolute goals	.19 (85%)	.03 (15%)	.19 (76%)	.06 (24%)	.09 (42%)	.13 (58%)	.48
test score	.04 (29%)	.10 (71%)	.04 (25%)	.12 (75%)	.00 (0%)	.11 (100%)	.42
recommend.	.08 (31%)	.18 (69%)	.08 (28%)	.21 (72%)	.04 (17%)	.20 (83%)	.51

Conclusions

It was our aim to better define the factors that mediate the influence of SES on learning achievement and the principal's advice at the end of primary education. We have shown that, taking into account differences between children with respect to aptitude and previous learning achievement, parents' expectations mediate the effects of socio-economic status on teachers' aspiration levels, test scores and recommendation at the end of primary education. This effect pertains to the preferred choice of secondary education for those parents who have expressed an explicit preference. The influence of parents' expectations is not restricted to the school careers of their children, but has also an impact on teachers' aspiration levels for their children. One could say that aspiration levels set by parents work out as a form of 'educational effectiveness'.

Figure 1
Model for aspiration levels, learning achievement and recommendation

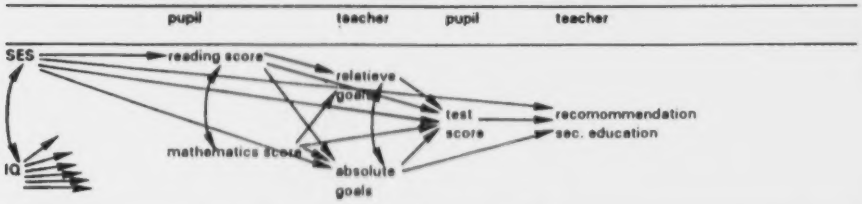
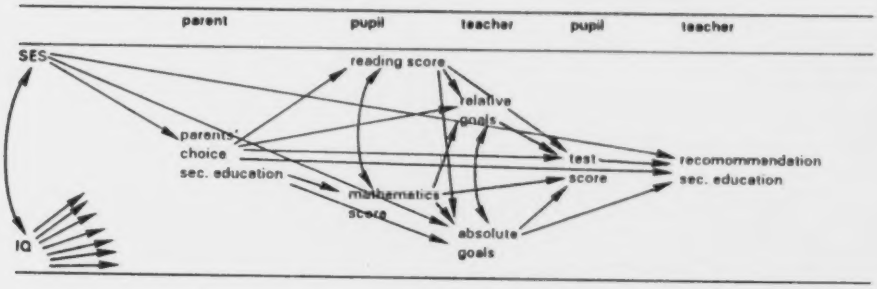


Figure 2
Model for aspiration levels, learning achievement and recommendation



1307

ACHIEVEMENT AND ACHIEVEMENT MOTIVATION

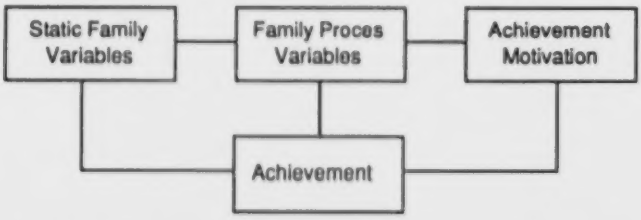
Dilek Ardaç, Turkey

Turkey bears a high demand for education and limited educational resources. As a natural consequence of these features, Turkish educational system can be characterized as being highly selective. Two basic milestones of selection are the entrance exams given before secondary school level and the University Entrance Exam. Acceptance to an outstanding college or highschool at secondary level is considered to be a favourable situation, because 1) it partially guarantees success at the University Entrance Exam; 2) most of these outstanding schools at secondary level offer their courses in a foreign language, which proves to be a highly valued asset for future vocational opportunities. Therefore preparation for the entrance exams given before secondary school level constitutes a major task in the lives of primary school students and their parents. Majority of parents consider it to be an important period, and seek professional advice in order to improve academic achievement of their children. Present study was an attempt to answer parental educational needs in this area. The purpose of the study was to identify parental factors related to achievement motivation and achievement of children in order to explain how parental factors act so as to influence academic achievement.

A total of 1420 primary school students and their parents participated in the study. The sample was chosen among subjects who were getting prepared for the entrance exam at the secondary school level. Therefore all the students and parents in the sample were facing a crucial achievement situation. This was considered to be a favourable fact in highlightening their achievement oriented behaviour so as to make it more easily detectable. The students in the sample were studied in terms of two main attributes: 1) Academic achievement; and 2) Achievement Motivation. Parents in the sample were investigated in terms of: 1) Home Educational Environment; 2) Parent Socialization; and 3) Parental Attitudes towards Child Rearing. Measuring instruments used to assess achievement and achievement motivation of children were: 1) Achievement Motivation Scale and 2) Achievement Tests given as a part of the private preparatory institute. A total of 46 formative achievement tests were included in the analysis. The instruments used to assess parental variables were: 1) Home Educational Environment Questionnaire, 2) Parent Socialization Questionnaires, and 3) Parental Attitudes Research Instrument. Home Educational Environment Questionnaire consisted of items concerned with parental behaviour and attitudes related to their children's educational activities such as the time and money spent on education, the degree of control over educational activities, and educational aspirations. Parent

Socialization Questionnaires consisted of items related to parental behaviour concerning the nature and time of demands and restrictions imposed on the child; and the kind of rewards or punishments provided to the child in response to actualization of demands and violation of punishments. The validity of the questionnaires were tested using: 1) Factor Analysis and 2) Results of previous correlational studies. Reliability of the questionnaires were determined by: 1) Item Analysis and 2) Reliability coefficients.
The analysis of data was carried out within the framework of a model based on empirical and theoretical evidence integrating parental variables related to the development of achievement motivation and achievement (Fig.1). Sociocultural characteristics of parents and family structure were considered as static family variables whereas parent socialization, child rearing practices and home educational environment were considered as family process variables. The data was analyzed using: 1) Pearson Product Moment Correlations; and 2) Analysis of Variance Technique.

Figure 1
Model integrating parental variables related to the development of achievement motivation and achievement of children



The results of the study indicate the existence of global and specific parental behaviours/attitudes related to different achievement levels of children. When significant correlations between the variables in the model were considered, it was possible to represent the global characteristics of the families whose children achieve higher compared to other children by the following schema:

Static FamilyCharacteristics:	Family ProcessCharacteristics:
Higher educated parents	More democratic attitudes
Parents with a degree in scientific or technical fields	Less overprotective attitudes
	Less disciplinarian attitudes
	Earlier demand of independent behaviour
	Preference for mild punishments
	Higher educational aspirations
	Higher value placed on education

1309

However the relationships between family variables and achievement motivation were low and no significant relationship was observed between achievement and achievement motivation of children. Possible explanations for low correlations observed between achievement motivation and parental variables as well as the nonexistence of a relationship between achievement motivation and achievement of children were discussed in terms of:

- the relative position of achievement motivation within the sample population in terms of achievement motivation scores with reference to crosscultural differences;
- the dominant effect of ability over achievement motivation in explaining the variance in achievement;
- homogeneity of sample population in terms of achievement motivation scores;
- problems concerning the content validity of the Achievement Motivation Test..

The implications of the results were mainly derived from parental factors showing significant correlations with academic achievement of children. Characteristics of parents whose children achieve higher compared to those with lower academic achievement were outlined resulting in a list of "favourable" parental attitudes and behaviours related to high academic achievement.

References

Atkinson, J. W. (1958). Motives, fantasy and action in society. New Jersey: D. Van Company

Alwin D. F. ; Thorton, A. (1984). Family origins and the schooling process: Early vs. late influences of parental characteristics. American Sociological Review, 49, 784-832

Bolles, R. J. (1975). Theory of motivation. New York: Harper and Row Publishers

Elliot, E. S. and Dweck, C. S. (1988). An approach to motivation and achievement. Journal of Personality and Social Psychology, 54(1), 5-12

Marjoribanks, K. (1987). Ability and attitude correlates of academic achievement: Family group differences. Journal of Educational Psychology, 79(2), 171-178

Walberg, H. J. and Marjoribanks, K. (1976). Family environment and cognitive development: Twelve analytical models. Review of Educational Research, 46, 527-551

Wood, J.; Chapin, K. and Hanich, M. E. (1988). Family environment and its relationship to underachievement. Adolescence, 123, 283-290

THE MEANING OF THE OCCUPATIONS OF WORKING MOTHERS FOR THE EDUCATIONAL INEQUALITY

J. Dronkers, SCO, University of Amsterdam, Amsterdam, The Netherlands

Working mothers with children of school age are a recent phenomenon in the Netherlands, but are more common in other European societies. The social and scientific meaning of this work for life chances generally and for educational inequality specially is still not clear and subject of heated debate (Goldthorpe, 1983 & 1984; Heath & Britten, 1984; Marshall et all., 1988; Millne et all., 1986; Stanworth, 1984). In this paper the effects of paid work outside home by mother and of the level of her work on the educational attainment at the end of primary school are estimated with ANOVA and LISREL for the most recent nationally representative cohort of Dutch primary school-leavers of the late '80s (the so-called VOCL-'89 cohort), controlling for other characteristics of herself, household, husband and children. It is a follow-up of a comparable analysis of an older cohort of Dutch primary school-leavers of the late '70s (Dronkers, 1989). The results in this most recent analysis show that the dilemma between work outside home and housewife is a false one. The question for the educational is not whether a mother works or not, but the level of her work is the central question. Working in labour jobs has a negative effect on children's educational attainment compared to housewife, in contrast to the positive effects of working as shopkeeper, farmer or employee, also after controlling for all other relevant characteristics. These effects of the level of mothers job on the educational attainment of her children are substantial and cannot be explained by other characteristics, but they are not equal to the effects of the level of her husbands job on the educational attainment of their children. This means that the conventional view of Goldthorpe and the radical view of Stanworth must be replaced by a more moderate one. This means also that mothers job has to be included in any measurement of parental class of pupils in order to get a accurate estimation. The effects of the level of mothers job on the educational attainment of her children don't differ between boys and girls. This result doesn't neither confirm suggestions of the model-function of working mothers for their daughters nor confirm suggestions of the need of boys to be controlled by a mother at/home. The effects of the level of mothers job do also not differ between ethnic groups. The last two results are contrary to some results from the USA. In the paper the effects of mothers job on the children's educational attainment in a nationally representative cohort of the late '70 (the so-called SMVO cohort) are compared to those of this recent VOCL-'89 cohort. Despite the growth of the percentage of working mothers during the '70s and '80s in the Netherlands

1311

(mostly part-time jobs) these effects of the level of mothers job did not change. A possible consequence of these results is that a general change of the tax and wages system from one based on households with one wage-earner to one based on individuals might be detrimental to families of parents with a low educational level, who are constrained to less qualified jobs.

References

Dronkers, J. (1989). "Working mothers and the educational achievements of their children." pp. 185-198. In: The social world of adolescents: International perspectives, edited by K. Hurrelmann and U. Engel. Berlin / New York: Walter de Gruyter.

Goldthorpe, J. H. (1983). "Women and class analysis: in defence of the conventional view." Sociology 17:465-488.

(1984). "Women and class analysis: reply to the replies." Sociology 18:491-499.

Heath, A. &N. Britten (1984). "Women's jobs do make a difference: a reply to Goldthorpe." Sociology 18:475-490.

Marshall, G., H. Newby, D. Rose and C. Vogler. (1988). Social Class in Modern Britain. London: Hutchinson.

Milne, A., D. Myers, A. Rosenthal & A. Ginsburg, (1986). "Single Parents, Working Mothers and the Educational Achievement of Schoolchildren." Sociology of Education 59:125-140.

Stanworth, M. (1984). "Woman and Class Analysis. A Reply to John Goldthorpe." Sociology 18:159-169.

EQUALITY AND ACHIEVEMENT IN EDUCATION AFTER EDUCATIONAL EXPANSION

Heiner Meulemann, Sozialwissenschaftliches Institut, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany

Abstract

Two questions are examined: (1) Did educational expansion attain the proclaimed goal of equalizing opportunity? (2) Did it have consequences beyond this political goal which are still relevant for status attainment, i.e. of a structural as against a cultural type? As regards the first question, a review of the empirical literature gives a predominantly negative answer. As regards the second question, it is argued that consequences of educational expansion still relevant for status attainment must be sought in the domain of achievement. Achievement is as fundamental a value in education as equality; and it is not only a means of measurement within education, but also a means of communication in the social context of education. However, while educational expansion has been advocated as a means of attaining equality, its consequences for achievement have been rarely considered. One would expect educational expansion to increase the importance of achievement in the attainment process, since it destroys traditional career expectations and selection procedures, and forces teachers to base their transition advice more strongly on achievement. Educational expansion pushes the educational system from sponsored to contest mobility (Turner). To support this hypothesis, three arguments are given to show that educational expansion has increased the importance of achievement.

First, path analyses of the influence of the father's occupation and school grades on the transition into selective secondary education for periods before and after educational expansion in the Netherlands and in Germany show that the importance of grades has increased at the expense of the father's occupation. Although the total influence of social background on educational attainment has not been diminished, legitimate criteria of educational selection have been reinforced at the expense of illegitimate criteria. Second, even if educational expansion was unable to narrow the gap of educational opportunity between social classes in the access perspective, it ought to have improved the percentage of lower class children in selective secondary education in the recruitment perspective. The resulting greater heterogeneity of selective secondary education should then devaluate the cultural capital of higher social classes. Advantages of higher class students resting on a natural accord of their home culture with the culture of the school are then reduced. However, in abscence of specific data, this

1313

remains merely a plausible hypothesis. Third, recent longitudinal between-country comparisons show that the inequality of occupational opportunity decreases - in puzzling contrast to the constant inequality of educational opportunity. This conclusion completely contradicts the generalizations established some 20 years ago (i.e. that the decreasing inequality of educational opportunity goes along with a constant inequality of occupational opportunity) and needs an explanation. One explanation may be that educational certificates have gained importance in occupational attainment; this explanation is validated in longitudinal data which show a decreasing influence of the father's occupation and an increasing influence of the son's education on the son's occupation during the period of educational expansion.

In conclusion, the increasing complaints about "pressure to achieve" in school are cited as an indirect, attitudinal indicator of the increasing importance of achievement. This negative evaluation, predominant in public opinion, contradicts the positive evaluation given here which rests on the aforementioned structural consequences of educational expansion. However, some tentative German evidence shows that this complaint is more prevalent in students from higher strata - which was to be expected if these students had to defend their opportunities during educational expansion by relying on achievement. Obviously, there are attitudinal costs for the structural consequences of educational expansion. Nevertheless, a positive evaluation of the increasing importance of achievement can be defended against the negative evaluation in public opinion.

SOCIAL STRATIFICATION, POLITICAL ATTACHMENT AND EDUCATIONAL EXPECTATIONS IN HUNGARY: AN ANALYSIS BASED ON PUBLIC OPINION SURVEY DATA

G. Halász, National Institute of Public Education, Hungary

The paper presents the main conclusions of two public opinion surveys that have been conducted recently in Hungary in order to describe education policy expectations of Hungarian society.

The survey have been conducted on samples representing the whole adult population of the country. The questions used in the questionnaires were linked with some crucial issues of educational policy like development priorities, pedagogic style of schools, family-school relationship, selection processes, the structure of the school system, local and central governance, attitude toward teachers, religious and moral education etc.. The education policy expectations and attitudes of social groups with different family background,m social status and political attachment have been compared.

The main conclusion based on the research can be summed up as follows:

- Policy expectations and attitudes concerning education in Hungarian society seem to be relatively well structured and stable as they are closely linked with both social status and political preferences.
- The most important conflicts that may appear in the educational policy arena are those between traditonal-collective and modern-individual task-definitions for schools, and between selective and egalitarian structural arrangements for the system.

1315

MINORITY LANGUAGE EDUCATIONAL RIGHTS: A CANADIAN EXAMPLE

Jonathan L. Black-Branch, Canada

Abstract

This paper examines the impact of the constitutional entrenchment of "minority language educational rights" on educational policy in Canada. Recent court rulings expanded the rights of minorities to include: the right to publicly funded minority language instruction; separate educational facilities from the majority language group; the school as a linguistic and cultural milieu for the minority language group; and educational governance by elected minority language officials. Specifically, this paper is based upon a study which developed analyses of court cases brought forward under the provisions of Section 23 of the *Canadian Charter of Rights and Freedoms.*

Introduction

Minority language educational rights has long been an important issue in Canada. To what degree should these rights be guaranteed and protected? English and French Canadians presently live in an officially bilingual and bicultural society. In 1982, Canada patriated its constitution becoming completely independent from Britain. The signing of the *Canada Act, 1982* marked this occasion. The Act includes the *Canadian Charter of Rights and Freedoms.* The Charter is the "supreme" law of the land. Canadians are now "guaranteed" basic rights and fundamental freedoms under constitutional law. Not only does the Charter reaffirm the dual languages of Canada, but Section 23 of the Charter guarantees "minority language educational rights" to both language groups. That is to say, English Canadians are guaranteed educational rights in areas where French is the the majority language group and French Canadians are guaranteed education in their language of instruction where English is the majority group.

Statement Of The Problem

This paper examines the impact of the constitutional entrenchment of minority language educational rights in Canada. Specifically, this study examined how major court rulings guaranteeing minority language educational rights affects educational policy. In particular, this study sought to answer: What impact has the constitutional entrenchment of minority language educational rights had on educational policy in Canada?

Research Methodology

Specifically, the paper is based upon a study which developed analyses of three court cases brought forward under the provisions of Section 23 of the *Canadian Charter of Rights and Freedoms*. This study employed two legal research approaches: conventional and fundamental.

- Conventional Legal Research focuses on court cases. Primarily, this approach to research revealed that which the courts have decided (judicial rulings) with regard to minority language educational rights in Canada. More specifically, the conventional research was "designed to collect and organise legal data, to expound legal rules" (Consultative Group on Research and Education in Law - CGREL, 1983, p. 65). The researcher identified three "major" court cases regarding minority language educational rights, under Section 23 of the Charter of Rights and Freedoms. Subsequently, the analysis focused on the written judgments in each case and reasons stated for such court decisions.
- Fundamental Legal Research rejects the use of "code", "case" and "statute" as the exclusive subject of analysis, and focuses on the impact of court decisions. This approach revealed what specific decisions, and the Charter in general, means for the administration and governance of education in Canada. In general terms, the purpose of fundamental research was to secure an understanding of the law, and its implications for school systems and minority language rightholders in Canada. It was "designed to secure a deeper understanding of law as a social phenomenon, including . . . historical, philosophical, linguistic, economic, social or political implications of law" (CGREL, 1983, p.66). Specifically, the purpose of this approach was to consider the present and future implications of these major court rulings and their effect on education in Canada.

Findings

The Canadian judiciary are upholding minority language educational rights as guaranteed under Section 23 of the Charter of Rights and Freedoms. Canadian courts have interpreted minority language educational rights in favour of the minority language groups. With these court interpretations comes the expansion of minority language educational rights in instances where school systems did not fully acknowledge such rights. For example, recent decisions by the Supreme Court of Canada and other superior courts, guarantee, where numbers warrant: the right to publicly funded minority language instruction; separate educational facilities from the majority language group; the school as a linguistic and cultural milieu for the minority language group; and educational governance by elected minority language officials. As a result, changes in legislation and educational policy have evolved.

Conclusion

These major court decisions have had a profound effect on the administration and governance of education throughout Canada. Major changes in education have evolved including an expansion of power and a reallocation of resources to the minority group. This power serves to alter the status quo of the majority language group.

1317

Discussion

At least three indicators support the educational and research value of this study. Firstly, this study revealed that minority language educational rights are guaranteed within the Canadian democracy. That is to say, the courts have upheld such constitutional rights. Secondly, this study addressed the perceived interests of academics, legal practitioners, educational administrators, elected school trustees, and education interest and human rights groups. These groups require extensive information regarding the constitutional entrenchment of minority language educational rights, and the implications of such rights for schools and school systems in Canada. Thirdly, the results of this study may provide a basis for comparison with other countries where minority language populations are presently seeking educational rights.

THE SCHOOL CAREERS OF A THIRD-GENERATION ETHNIC GROUP: MOLUKKAN STUDENTS IN THE NETHERLANDS

M.C. Robijns, J. Roeleveld, & A.M. Veen, Stichting Centrum voor Onderwijsonderzoek (SCO), University of Amsterdam, Amsterdam, The Netherlands

Introduction

The school careers of pupils belonging to an ethnic minority group are generally far from acceptable compared with those of native Dutch pupils. It is however often assumed that the longer the stay in the host country, the better the position in terms of school career. According to this view, the educational results of a later generation will be more in line with those of the dominant group than those of the preceding generations, the situation thus improving step by step. The reason is that the second- and third generation of people belonging to a minority group have a better understanding of the language and the culture prevailing in the host country and are therefore better equipped with skills and tools enabling them to function in the new society and compete with their 'indigenous' fellowmen.

The data we possess on one of the ethnic minority groups in the Netherlands, the Moluccans, the first ethnic minority with an ancienity of three generations, show however that the 'stay duration' alone isn't enough to improve the school careers of pupils belonging to a minority group.

At this moment, there is a third generation of Moluccan pupils populating the Dutch secondary schools. Research data yet proved that Moluccans still hold a lower position compared to their Dutch fellowmen in terms of the results obtained in school and on the labour market: a high rate of grade repeating, high dropout- and unemployment rate and an exceedingly large presence in lower forms of secondary education.

People from the Moluccas, (the south-east of Indonesia), are one of the first ethnic groups that came to live in the Netherlands. They live here since Indonesia, formerly a Dutch colony, obtained its independence. They arrived in 1951 as a group of ± 12.500 people. Compared to most of the other ethnic groups in the Netherlands, their migration motive was not labour but (for most of them) the fact that they had serviced in the Dutch army before the independence was declared. Therefore, it was no longer possible for them to stay in Indonesia as long as there was no independent Moluccan state. Although initially a temporary stay was intended, gradually it became clear that a permanent stay of the Moluccans in the

13i9

Netherlands had to be taken into consideration. The chance of success in establishing the - much desired - independant Moluccan Republic in the Indonesian archipelago turned out to be a less and less realistic option and, with this, their possibility to return vanished.
Due to this special historical, colonial relationship with the Dutch, Moluccans differ in several respects from other ethnic groups in the Dutch society: their occupational background (originally military) and their housing situation for instance (Moluccans were initially accommodated 'en groupe' in areas with few possibilities on the labour market).
Despite the differences between their group and other ethnic minorities living in the Netherlands, the research data concerning Moluccans can still give a general indication on the educational position held by ethnic minorities after a stay of several generations in our country.
The Centre for Educational Research (SCO) conducted a research concer-ning the development and background of the school careers followed by Moluccan pupils in the Netherlands. The research has been financed by the Dutch Institute for Educational Research (SVO).
In our paper we shall present part of the results we obtained in this research.

Method
A matching procedure has been applied in order to compare the school careers of Moluccan and Dutch pupils. The data were provided by the administration authorities belonging to secondary schools with a concentration of Moluccan students (>10).
The research was focussed on pupils who entered school in 1987-1988 (cohort '87-'88). Pupils belonging to this cohort were matched on gender, father's occupation, age and educational advice. For the reconstruction of their school careers, data were collected on the schools we mentioned above, concerning their educational position during successive school years ('87-'91), school type, dropout rate and -reason. Other (background) information was gathered by means of questionnaires addressed to students and their parents containing questions on language orientation, school experiences, self-image, aspirations and expectations with regard to future education and occupation.

Results
In order to bring a better understanding in the school careers of the Moluccan pupils, we shall present a school-career model, relating individual student characteristics and results in the secondary education: a model containing data on 123 Moluccan pupils in which the directions of Moluccan school careers will be displayed.
Subsequently, data on 69 match-pairs will be analyzed in order to discover differences in the school careers of Moluccan and Dutch pupils. The models contain background variables (gender, socio-economic status) as well as variables indicating the direction of their school careers (including primary education), position held in secondary education and variables like language orientation, experience, aspirations and future expectations. By using the matching procedure, differences due to socio-economic and gender factors will be excluded. For the analysis, a Lisrel procedure was used.

Conclusions

Regarding school careers, significant differences have been found between Moluccan and Dutch pupils mainly in point of "grade repeat-ing". Furthermore, we found differences in pupils' "attitude towards school" and their"learning perception".

Within the Moluccan group, differences have been found in the school-career-model related to the cultural orientation of the pupils (as measured by a combination of data concerning "language-orientation" and "within-group-relationships") and, also, their housing situation (whether or not living in a Moluccan neighbourhood -"wijk"-).

OP-STAP, A HOME INTERVENTION PROGRAMME PREPARING CHILDREN FOR SCHOOL: ITS EFFECTS

P. Vedder, & L. Eldering, Centre for Intercultural Pedagogics, Leiden, The Netherlands

In 1987 the Dutch Ministry of Welfare decided to start a study on the possibility of using an Israelian home intervention programme, called HIPPY, in the Netherlands. HIPPY aims at improving the 4 to 6 year old children's preparation for school. In the Netherlands it is mainly used with immigrant families. The study's goal was twofold: We had to analyze the feasibility of implementation and we had to assess the effects. This paper deals with the effects.
Participating children get a structured curriculum aimed at the development of a broad range of cognitive and linguistic skills. The programme also aims at a change of the mothers' educational practices. They should experience that they can be effective teachers and thus become the agents of a developmentally more stimulating environment for the children.
We used a quasi-experimental group design with pre- and post testing. Experimental and control group were matched on the variables neighborhood, ses, ethnicity and ses. Children's intelligence was pre and post tested. After finishing the programme children's language skills were measured and teachers rated the children's class behavior. Only mothers participating in the programme, the experimental mothers, were interviewed twice before starting the programme. The same was done after finishing the programme. We measured the cognitive home stimulation, educational expectations and educational practices. A total of 141 families participated in the programme. Hundred and five children were in the control group.
We found some cognitive and linguistic effects with the children. Information on the implementation of the programme is used as a possible explanation for the findings. Actually, the extend of participation appeared to be positively correlated with effects. An important problem seems to be that overall participation was not intensive. We discuss implications for programme adaptation and implementation strategies.

OP-STAP, A HOME INTERVENTION PROGRAMME PREPARING CHILDREN FOR SCHOOL: ITS IMPLEMENTATION

L. Eldering, & P. Vedder, Centre for Intercultural Pedagogics, Leiden, The Netherlands

In 1987 the Dutch Ministry of Welfare decided to start a study on the possibility of using an Israelian home intervention programme, called HIPPY, in the Netherlands. HIPPY aims at improving the 4 to 6 year old children's preparation for school. In the Netherlands it is mainly used with immigrant families. The study's goal was twofold: We had to analyze the feasibility of implementation and we had to assess the effects. A total of 141 families participated in the programme. This paper deals with the implementation.

We describe strategies to reach families and to maintain them in the programme. We did a non response study. We collected some background information about all the mothers who were personally contacted for participation. Moreover, with about 100 mothers we had a lengthy interview on the home situation and educational practices. Afterwards, 45 of these mothers decided not to participate. The collected information could be additionally used in our non response study. We also registered the number of weeks families participated and the number of contacts families had with paraprofessional programme instructors. Drop-out families were interviewed shortly after the drop out had been reported in order to explore reasons for drop-out.

Generally it was difficult to reach families, although some strategies were more successful than others. In the non response analysis we found no clear reasons or family characteristics that explain why some families participated and others did not. In the drop-out study it appeared that changing family circumstances were the main reason for drop-out. Still, 40% of the reasons given by mothers referred to the programme. We had expected that families less acquainted with western culture would drop out more rapidly and more frequently than more western oriented families. Although it was difficult to reach the families characterized by a greater distance towards western lifestyle, once they participated, it appeared relatively easy to keep them in the programme. The intensity of participation was slightly disappointing. The families that completed the programme worked on the average with 42 of the 60 available week programmes, with a standard deviation of 10 weeks. We shall discuss implications of the study for future implementation attempts.

1323

PYGMALION AND EFFECTIVENESS OF 'BLACK' SCHOOLS: TEACHERS' STEREOTYPES AND COVERT TARGET LEVELLING TOWARDS MINORITY PUPILS

Paul Jungbluth, ITS, Nijmegen, The Netherlands

Summary

First aim of this research report is to identify determinants of educational achievement of migrant pupils in Dutch elementary schools. In the scheme of variables, teachers' perception of individual pupil's capacities and of attainment targets covertly levelled for each individual pupil are crucial. We surveyed 807 pupils and their class-teachers in 44 elementary schools with all at least 15% migrant pupils.

Indeed both kinds of teachers' perspectives significantly contribute to differences in achievement in relation to pupils' social background. The minority status or (supposed) ethnic identity of the pupils appears to be relevant mainly because of the social class identity that (as a matter of historical coincidence or economic determination) often goes with it.

Although this 'ethnic' variant of Pygmalion gives relevant clues at individual level as to what determines educational achievement in 'black' elementary schools, it does not fit to explain school related differences in effectiveness. At the individual level it only explains that part of 'ethnic' differences in achievement that can be rerooted to the parental social class.

Reason and research questions

The research described adds up to a series of four studies, which were subsidized by the Institute for Educational Research in the Hague (SVO; project number 8095). Central in this is the progression of migrants from primary to secondary education. By now, it is common knowledge that most migrant pupils in the Netherlands stay far behind the average indigenous students. In particular Turkish and Moroccan students significantly more often progress to the lowest forms of secondary education, than Dutch students do (cf. Tesser et al, 1989). Generally, when looking for an explanation, one emphasises the external characteristics of the migrants: different language, different culture, their length of stay in the Netherlands, eventually also their socio-economic background, and the educational support at home (cf. Van Langen & Jungbluth, 1990).

In this article we will focus on one possible school-related determinant of educational effectiveness for migrant pupils: the Pygmalion-complex. Our present research involves forty-four primary schools with a school population that

embodies a relatively large number of migrants. Three major questions will be answered:

- are teachers' perceptions of pupils' abilities ethnically biased and if so, does it affect the implicit goal levels teachers set for each individual pupil?
- to what extend can such ethnic aspects of the Pygmalion-complex explain ethnic inequality in educational achievement?
- what is the relation between teachers' social stereotypes or matching differentiation practices and the effectiveness of the schools?

Database

The data were collected at 44 primary schools in the spring of 1990. At the time the data were collected, all 44 schools had at least four migrants in group eight (age twelve), mainly Turks or Moroccans. Data were collected at different levels. The students of group eight took language and arithmetic tests, which consisted of a selection from CITO entrance tests (Moelands, 1988). In addition they took a non-verbal IQ-test, consisting of the section 'Draaien' (Turned figures) of the ISI series. The database contains information on 807 students, of which 434 are indigenous, 160 Turkish, 97 Moroccan, 45 Surinamese, 5 Antillean, 3 Mollucan and 63 others (varying from south and north-west Europeans to Africans and Asians). The percentage of migrants on the total school population at the 44 schools varies from 8 to 97%, with an average of 48%. This clearly shows that this is definitely not a national representative sample survey. Not only migrants are over-represented, but also Dutch working class children, who are usually over-represented in migrant classes.

Findings and discussion

The results from our analysis reveal that our intention has only partially been successful. Part of the original 'ethnic' differences in achievement can be seen as an outcome of the social class composition of certain minority groups in the Netherlands. Both teacher perceptions and target levelling towards migrant pupils seem to be social class related in as far as they are originally different for different ethnic groups: minority pupils (possibly most the ones with relatively good capacities) suffer from a social class bias in teacher behaviour. Only in considering this complex interaction, we do not conclude for racist bias in teacher perceptions, which could eventually be done from the original findings.
In last instance our analysis forces us to differentiate between educational effectiveness and school effectiveness. The processes we have tried to indicate, are of relevance in first instance, at the individual pupils' level. When it comes to explaining school differences, they seem to be of no value, eventually just as a suppressor variable.

DIFFERENCES IN POLICY MAKING CAPACITY AND CULTURE: AN INVESTIGATION IN TWO EDUCATIONAL PRIORITY AREAS

Inger Berings, Heidi van Cuijk, Peter van der Kley, & Anton Wesselingh, Department of Educational Sciences, Nijmegen University, Nijmegen, The Netherlands

Results of earlier studies with regard to the implementation of educational priority policy (EPP), have indicated that the implementation of an innovation in a professional bureaucracy (like schools and social welfare organisations) produces difficulties, due to specific structural and cultural characteristics of these professional bureaucracies.

In this study, that is based on several organisational theories, we have hypothetically determined structural and cultural conditions, that schools and organisations should realise within their organisations, in order to improve the implementation of EPP.

The emphasis in this study lies upon the structural element. However, as an organisational innovation cannot be achieved by structural changes alone, the culture of the organisation must change as well. For that reason the cultural element is included.

The structural conditions are realised when the school or organisation disposes of an adequate policy making capacity (PMC). The term policy making capacity is taken from the theory of E. Marx. PMC contains two aspects: it consists of the formation of policy, but at least as important is the execution of this policy. It is necessary that the policy, formulated by the management, is executed by the workers of the organisation.

Schools/organisations dispose of an adequate PMC in cases when:

- the organisation has a need to use policy as coordination mechanism;
- there is sufficient knowledge of the external environment;
- the organisation has a good equipment for policy making at it's disposal;
- the authority structure contributes to an adequate policy making;
- the management is able to direct the teachers/workers by means of steering mechanisms;
- the management is able to stimulate the teachers/workers to conform themselves to the innovation.

The cultural conditions are realised, when there is a certain extent of internal consistency of opinions and perspectives on education and learning in general, and more specific in the relation to EPP.

The purpose of this study is to make a comparison of schools and organisations in two different educational priority areas (EPA's) in terms of their PMC and their cultural climate.
Our study is based upon surveydata, obtained by personel (teachers and welfare workers) actually engaged in EPP-activities (N=250). The data are analysed with the help of the statistical program SPSS-PC. Several statistical methods have been used, such as factor-analyses, breakdown-analyses and disciminant-analyses.
In the first place we examine how far the schools and organisations in two educational priority areas (Nijmegen/Groesbeek and Tiel), have realised the structural and cultural conditions that improve the implementation of EPP. In other words, do these schools and organisations dispose of an adequate PMC and is there a certain extent of consensus and internal consistency?
The results indicate that there are several significant differences between the two educational priority areas, both in structural and cultural aspects of PMC. *Structural* conditions that display a significant discrepancy are 'knowledge of the external environment' (in particular knowledge of the areaplan) and the 'role of the internal EPP-coordinator to induce cohesion in EPP-activities inside schools and welfare-organisations'. The important *cultural* conditions that differentiate between the two EPA's are concerned with 'orientation towards education, learning and schooling' (progressive versus traditional) and 'opinions about cooperation between welfare-organisations and schools' (positive versus negative).
In the EPA Nijmegen/Groesbeek the respondents are not as well informed of the areaplan as their colleagues in the EPA Tiel; on the other hand the EPP-coordinators of welfare-organisations or schools are more active in establishing cohesion in EPP-activities than their colleagues in Tiel.
In Nijmegen the dominant orientation towards education, learning and schooling tends in a progressive direction, while in Tiel the respondents tend to be more traditionally orientated. The attitudes towards cooperation between welfare-organisations and schools are more positive in Nijmegen than in Tiel.
The results can be interpretated in terms of the differences in history, tradition and development of both EPA's. The establishment of educational priority policy in the early 80's in Tiel has created a more coherent whole of activities that aimed at reducing inequality in educational processes and results of girls, lower class students and students of etnic minority groups than before. In Nijmegen, however, efforts of schools to reduce inequality in education date back to more than twenty years. The development of educational priority activities there can be labeled as a 'bottum-up' strategy: schools, lateron in cooperation with welfare-organisations, developed a policy themselves. Against this background, the areaplan plays a different and less important role than it does in Tiel. Here the areaplan was the starting point of EPP-activities and therefore it can be labeled as an 'top-down' approach; in Nijmegen it functions as the final formalization for bureaucratic purposes of already existing activities.
These differences in history and experience with EPP-activities between the two EPA's have had an impact on the role of the EPA-coordinator: in Tiel his presence is more directive and stimulating. In Nijmegen, on the contrary, the role of the internal coordinator of the participating organisations and schools is more important than the role of the EPA-coordinator. In Tiel the organisations (by their lack of experience) have a need for more coordination from outside the

organisation. This fact can explain the second structural difference between the two EPA's.
These differences have had their consequences for the organisation structure of the areas themselves. In Nijmegen the EPA is divided in several smaller regional areas; in Tiel there is a division in sectors (primary-, secondary education and welfare).
The culture of the two EPA's is influenced by the differences in structure of the areas. A division in regional areas, like in Nijmegen, leads to a more intensive cooperation between schools and welfare-organisations and the corresponding attitudes towards such a cooperation. In Tiel it is difficult to cross the borders of the sectors, because they have a preference to attack inequality in education within their own sector.
Nijmegen on the other hand has a preference for a sector-exceeding, but regionbounded approach. These preferences are related to (differences in) orientations towards education, learning, schooling and educational priority policy.
Summarizing, our results suggest that in EPA Nijmegen, in comparison with EPA Tiel, there have been developed better conditions for an effective and successful implementation of EPP. Although we cannot determine the effectivity of reducing educational inequality by these two EPA's on a pupil level, we can reach to a conclusion on a *conditional* level.

References

Marx, E.C.H. (1987). Vermogen van scholen tot het voeren van bestuurlijk beleid. In: Genemans, L. (red.) Autonomie in scholen en deiregulering, Nijmegen, p. 7-27.

Marx, E.C.H. (1988). De uitkomsten van de OTO-gevalsstudies, organisatiekundig bezien. In: MESO, 1, 20-32.

ACHIEVEMENTS IN PROBABILITY OF STUDENTS FROM OPPSITE SIDES OF THE WORLD

B.R. Zinn, & Y. Voronel, Rothberg School for Overseas Students, The Hebrew University of Jerusalem, Mount Scopus, Jerusalem, Israel

Introduction

The experience of many college faculty members in education and in social sciences indicates that a large proportion of university students in introductory statistics courses do not understand many of the concepts that they are studying (Garfield and Algren,1988). Students often tend to respond to problems by "number crunching", plugging quantities into a computational formula or procedure without forming an internal representation of the problem.Students appear to have difficulties developing correct intuitions about fundamental notions of probability as many students have an underlying difficulty with rational number concepts and proportional reasoning, which are used in calculating, reporting and interpreting probabilities. Some difficulties, especially with conditional probability, seem to be caused by formal wording or algebraic notation (Pollatsek et al.,1987).

This study examines the acquisition of probabilistic thinking skills of two student groups in a pre-academic program for overseas students at the Hebrew University: Spanish-speakers from Latin America and Russian-speakers from the former USSR. We postulate that difficulties experienced by these two student groups will have different sources. For the Spanish-speakers, it may be based on didifficulties with proportional reasoning and algebraic manipulations. For the Russian-speakers, there may be a lack of flexibility in moving from one representation to another of the same data. Do different education systems produce different thinking patterns?

Method

Target population

Students opting for the humanities or social sciences trend from these two sub-populations of Spanish (33) and Russian (174) speakers formed our sample. All students were exposed to an introductory course in probability.

1329

Instruments

- Placement Test
  All entering students took a placement test in mathematics. This test comprised 70 multiple choice items covering knowledge in several areas, namely: arithmetic, algebra, exponents, trigonometry, logarithms, functions and probability.
- Achievement test in probability
  This test was administered to all students at the end of a semester course dealing with basic notions of probability. It consisted of 52 multiple choice items covering such notions as addition and multiplication rules for probabilities, independent and mutually exclusive events, conditional probability, Bayes rule, random variables, binomial and hypergeometric distribution functions.
- Questionnaire evaluating course
  This questionnaire was answered anoymously by students towards the end of the semester. It consisted of 6 items dealing with the amount of knowledge acquired, the organisational aspects at course level and at the individual lesson level, the interest generated, the contribution of exercises,as well as an overall evaluation of the course. Each item was scored on a scale from 1 to a maximum of 20.

Results and discussion

The placement test showed that Russian-speakers had a better mathematical knowledge base and more algebraic skills than Spanish-speakers on entry. Both groups had little prior knowledge about probability before the course as shown in the table below.

Table 1

Entering mathematics knowledge (in %) for Spanish and Russian speakers

placement	mean $_{Spanish}$	mean $_{Russian}$
part 1	25.58	51.74
part 2	16.75	19.80

where part 1 deals with arithmetic, algebra, exponents, functions, logarithms and trigonometry and part 2 deals with probability and combinatorics.

Both groups were exposed to the same introductory course (90 hours) with the same exercises translated into Russian or Spanish. The only difference being that each group was taught in its own home language. The two groups scored similarly on the achievement test ($X_R = 74,6$; $X_{Sp} = 72,9$) contrary to our previous experiences where Russian-speakers had lost their initial advantage and scored considerably less than Spanish-speakers on open-ended questions rather than multiple choice items.

Table 2
Students' Achievements in %

content areas and skills	Spanish speakers (N = 34)	Russian speakers (N = 174)
different representations	84.4	72.5
sets	76.5	79.9
probability distributions	74.1	73.6
Bayes	69.3	70.9
random variable	67.4	65.1
independence	63.3	50.3
functions	61.2	82.8
algebraic manipulations	55.2	76.3
combinatorics	31.8	39.4

Spanish-speakers, as could be expected from their poor entry skills, had difficulties with items requiring algebraic manipulations but were able to move from one representation of data to another with ease. This was not the case for the Russian-speakers, who appeared to rely on the formulas and "number crunching". Both groups experienced conceptual difficulties with items dealing with combinatorics and independence. There were also problems with the language of probability such as "at least".

Table 3
Students' evaluation of course

	Spanish speakers (N = 27)	Russian speakers (N = 112)
knowledge	15.13	15.88
contribution of exercises	16.28	15.72
organisation of course	16.96	16.51
organisation of lessons	17.60	16.62
interest	17.73	16.46
overall evaluation of course	15.56	14.96

Both groups expressed satisfaction with the course, with the exercises and with the knowledge acquired. They felt that the course was well organised and interesting.

Conclusion
It appears that these two sub-populations do have different approaches to solving problems dealing with probability. The Spanish-speakers, despite their poor mathematical skills, do manage to form visual representations of the problems. The Russian-speakers seem to rely too heavily on formulas and problems appearing in a particular mode. These impressions merit futher investigations using interviews.

References

Garfield, J. and Algren, A.1988.Difficulties in learning basic concepts in probability and statistics : implications for research. Journal for Research in Mathematics Education 19 (1) 44-63.

Pollatsek, A., Well, A. D.,Konold, C., Hardiman, P., and Cobb, G. 1987. Understanding conditional probabilities. Organisational Behaviour and Human Decision Processes 40, 255-269.

EARNINGS, JOBS AND EDUCATION IN SPAIN

Rosario Martinez, José-Ginés Mora, & Luis Vila, Department of Applied Economics, University of Valencia, Spain

The Human Capital theory provides the usual framework for studying the relationship between earnings and education. Despite the fact that this theory has undeniable shortcomings, it is an appropriate tool as a first approach for analysing the effect that different curricula have on the wages earned during the working life. In this poster we present some results of an analysis on earnings and education in Spain, using the classical Mincer model. We have used data from the 1981 Family Budget Survey (a big survey gathering data of 24.000 families). We have only used householder s data since it was the only group about whom we had information about the level of education. In addition, we have eliminated from the sample all the householders self-employed and the employees in the agricultural sector. Firstly, we have analysed the main features of the data splitting for male and female, people younger and older than 45 years, emplo yed and unemployed, people living in cities smaller and larger than 50.000 inhabitants, and people living in North-eastern Spain and in the south-west regions of Spain. In table 1 and 2 we present the main results of this analysis.

Table 1

Number and percentages of people in different educational levels classified by gender and age.

education level	total	males	females	<45 yrs	>45 yrs
nostudies	2.287	2.107	180	922	1.365
	20%	92%	8%	40%	60%
primary 5 years	5.837	5.528	309	3.145	2.692
	50%	95%	5%	54%	46%
primary 8 years	1.161	1.056	105	812	349
	10%	91%	9%	70%	30%
secondary	1.113	1.029	84	761	352
	10%	2%	8%	68%	32%
higher education	1.213	1.082	131	753	460
	10%	89%	11%	62%	38%
total	11.611	10.802	809	6.393	5.218
	100%	93%	7%	55%	45%

1333

We can observe in these tables that the level of unemployment decreases with the level of education. On the other hand, younger peop le and people living in large cities have a higher level of education. There are not important differences between the educational l evels in North-eastern Spain (the richest zone) and the South-west region, with the exception of the people without studies, that are more frequent in the poorer regions. We can not obtain definite conclusions about differences between men and women because the fe male householders are not representative of the average woman in Spain.

Table 2
Number and percentages of people in different educational levels classified by employment status and location

education level	total	employ	unemp	small cities	large cities	north east	south west
no studies	2.287	1.948	339	1.108	1.179	695	1.592
	10%	85%	15%	48%	52%	30%	70%
primary 5 years	5.837	5.436	401	2.008	3.829	3.051	2.836
	50%	93%	7%	34%	66%	52%	48%
primary 8 years	1.161	1.120	41	284	877	539	622
	10%	96%	4%	24%	76%	46%	54%
secondary	1.113	1.048	29	191	922	588	525
	10%	94%	6%	17%	83%	53%	47%
higher education	1.213	1.194	19	232	981	569	644
	10%	98%	2%	19%	81%	47%	53%
total	11.611	10.782	829	3.823	7.788	5.392	6.219
	100%	93%	7%	33%	67%	46%	54%

In table 3 we present the level of earnings for different educational levels, taking 100 as a base for the average earnings of the s ample. As was expected, the earnings increase with the educational le vel. Wage earning people with a higher education degree earn 65 % more than the average and more than double than people without studies. The most experienced people do not earn more than the less experienced on average, but, by educational level, this is only true for the less educated. The qualified people earn more when the y have more experience.

Table 3
Earnings by educational level for gender and experiance

education level	total	males	females	<30 years	>30 years
no studies	72	72	53	75	70
primary 5 years	89	91	67	88	90
primary 8 years	112	114	85	107	123
secondary	130	132	105	124	150
higher education	165	170	118	165	165
average	100	101	79	105	94

In tables 4 and 5 we present the results of the Mincer model for obtaining the return rate of the different educational levels The percentages showed in the first

part of each column of the tables mean the return rate for each level over the former. The second per centage in each cell represents the accumulated return rate over the people without studies.

Table 4

Earnings return rates by different educational levels for gender and employed people.

	total		males		females		employed	
	incr	acum	incr	acum	incr	acum	incr	acum
primary 5 years	17%		15%		40%		20%	
primary 8 years	25%	42%	25%	40%	33%	73%	26%	46%
secondary	17%	59%	16%	56%	23%*	96%	17%	63%
higher education	25%	84%	27%	83%	14%*	110%	27%	90%

* No statistically significant

Table 5

Earnings return rates by different educational levels for age and location

	<45 years		>45 years		small cities		large cities	
	incr	acum	incr	acum	incr	acum	incr	acum
primary 5 years	12%		21%		7%*		22%	
primary 8 years	25%	37%	28%	49%	25%	32%	23%	45%
secondary	17%	54%	19%	68%	16%*	48%	17%	62%
higher education	33%	87%	12%	80%	7%	55%	30%	92%

* No statistically significant

We can observe that the return from education is specially high for people with 8 years of primary education (25%) and for people with a university degree (25%). Despite the fact of women have a lower level of wages than men, the wage rates for women are higher, s pecially in the first educational levels. If we only select the employed people the rates are obviously higher than for the whole po pulation. The accumulated wage rate is higher for younger people, this is specially true of people with a university degree (curiou sly the generation with a greater proportion of university graduates). Lastly, it is obvious that rates of return are considerably higher for people living in cities than for people living in small towns or rural areas.

ENHANCING EDUCATIONAL OPPORTUNITIES

Greetje van der Werf, Institute for Educational Research (RION), Groningen, The Netherlands

Introduction

In the Netherlands, as well as in the most other countries in the Western world, there exists a strong relationship between socio–cultural and ethnic background characteristics of students and their learning outcomes and educational careers in primary and secondary education. Since 1985 the Dutch Ministry of Education carries out the Educational Priority Program in order to enhance the educational opportunities of children from lower income families and from ethnic minorities. This improvement should be realized by individual schools or by co–operations of schools and welfare institutions. The emphasis of the program is laid on the improvement of educational achievements in arithmetics and Dutch language in primary education.

The Priority Policy consists of two components, the general facilities component and the so–called area component. The general facilities component holds that all primary schools where children form the target groups are present receive extra teaching staff, proportional to the rate of children from lower income families and from families that belong to an ethnic minority group. Therefore all children are weighted. Dutch low SES children count for 1.25, immigrant children for 1.90. Also pupils who live in boarding institutions or whose parents are travelling around are weighted, with respectively 1.40 and 1.70. Pupils who do not belong to one of the mentioned groups are counted for 1.00. Per school the weighted number of children is divided by the absolute number and multiplied with 100. On the basis of the resulting score (the schoolscore), that can vary from 100 till 190, schools receive their teaching staff.

Another part of the program consists of providing priority areas with extra facilities. Such priority areas are consituted by schools in co–operation with educational and welfare agencies. To receive extra facilities the areas have tot submit a plan of activities to support schools and the other agencies in their priority area.

An important characterists of the Priority Program is that schools and areas are free to choose themselves their activities by means of which they want to enhance the educational opportunities of the target groups. This freedom probably will result in many differences between schools in priority characteristics on school and classroom level and in outcomes on the level of students. Some schools and areas will pay extra attention to the so–called effective school and instructional characteristics, in order to raise the quality of the whole schoolpopulation, others

will develop more specific strategies and programs for the disadvantaged target groups. There will be also schools and areas that will give extra attention to the parents of the target groups of the priority program.

In the symposium the focus of the papers lay on answering three main questions.

- What activities schools and priority areas are carrying out in order to diminish or to prevent educational disadvantages.
- Are there any differences between schools in educational outcomes and careers of the target groups of the priority program?
- To which characteristics of the school organization, characteristics of teachers and characteristics of the instructional processes the apparent differences can be ascribed.

Four of the five papers in the symposium are based on the data of the national evaluation of the priority program. This evaluation was started in 1987. The main part of the evaluation is a longitudinal cohort study in 700 primary schools. In this cohort study pupils in grade 4, 6 and 8 are tested every two years for aritmetic, Dutch language and intelligence. From the pupils in grade 8 also information about the recommendations for secondary education is collected. Data about social and ethnic background, about parental educational support and the use of cultural resources are collected from the parents. The data about school and instructional characteristies are gathered from school principals and teachers. The papers of Mulder, Weide en Van der Velden will report about specific analyses of these data in order to answer the questions about outputdifferences for the specific target groups of the priority program between schools and about the school and classroom characteristics to which these differences can be ascribed school. The paper of Weide focusses on ethnic minority children, the paper of Van der Velden on Dutch low SES children. Both papers analyse the output criteria in terms of students' achievements in Dutch language and arithmetics. The paper of Mulder focusses on the recommendations for secondary education for ethnic minority children in relationship with differences between schools in the proportion ethnic minorities within the school population. The paper of Jungbluth, although it is not based on the data of the evaluation study, also concentrates on the achievements of ethnic minority children, in his case children from Maroccan and Turkish origin. He relates differences in achievements to teachers' expectations about the capacities and background characteristics of these children.

In the evaluation program also some more qualitative studies are carried out. One of these is the study into the co–operation between the priority institutions (schools and areas) and basic adult education institutions. This co–operation seems to be important in order to enhance the educational level of the parents of the children from the target groups which in turn can have a positive effect on these children themselves.

In the paper of Reezigt the main conclusions of the existence of such a co–operation will be presented.

1337

PYGMALION AND EFFECTIVENESS OF 'BLACK' SCHOOLS; TEACHERS' STEREOTYPES AND COVERT TARGET LEVELLING TOWARDS MINORITY PUPILS

Paul Jungbluth, ITS, Nijmegen, The Netherlands

Summary

First aim of this research report is to identify determinants of educational achievement of migrant pupils in Dutch elementary schools. In the scheme of variables, teachers' perception of individual pupil's capacities and of attainment targets covertly levelled for each individual pupil are crucial. We surveyed 807 pupils and their class-teachers in 44 elementary schools with all at least 15% migrant pupils.

Indeed both kinds of teachers' perspectives significantly contribute to differences in achievement in relation to pupils' social background. The minority status or (supposed) ethnic identity of the pupils appears to be relevant mainly because of the social class identity that (as a matter of historical coincidence or economic determination) often goes with it.

Although this 'ethnic' variant of Pygmalion gives relevant clues at individual level as to what determines educational achievement in 'black' elementary schools, it does not fit to explain school related differences in effectiveness. At the individual level it only explains that part of 'ethnic' differences in achievement that can be rerooted to the parental social class.

Reason and research questions

The research described adds up to a series of four studies, which were subsidized by the Institute for Educational Research in the Hague (SVO; project number 8095). Central in this is the progression of migrants from primary to secondary education. By now, it is common knowledge that most migrant pupils in the Netherlands stay far behind the average indigenous students. In particular Turkish and Moroccan students significantly more often progress to the lowest forms of secondary education, than Dutch students do (cf. Tesser et al, 1989). Generally, when looking for an explanation, one emphasises the external characteristics of the migrants: different language, different culture, their length of stay in the Netherlands, eventually also their socio-economic background, and the educational support at home (cf. Van Langen & Jungbluth, 1990).

In this article we will focus on one possible school-related determinant of educational effectiveness for migrant pupils: the Pygmalion-complex. Our present research involves forty-four primary schools with a school population that embodies a relatively large number of migrants. Three major questions will be answered:

- are teachers' perceptions of pupils' abilities ethnically biased and if so, does it affect the implicit goal levels teachers set for each individual pupil?
- to what extend can such ethnic aspects of the Pygmalion-complex explain ethnic inequality in educational achievement?
- what is the relation between teachers' social stereotypes or matching differentiation practices and the effectiveness of the schools?

Database
The data were collected at 44 primary schools in the spring of 1990. At the time the data were collected, all 44 schools had at least four migrants in group eight (age twelve), mainly Turks or Moroccans. Data were collected at different levels. The students of group eight took language and arithmetic tests, which consisted of a selection from CITO entrance tests (Moelands, 1988). In addition they took a non-verbal IQ-test, consisting of the section 'Draaien' (Turned figures) of the ISI series. The database contains information on 807 students, of which 434 are indigenous, 160 Turkish, 97 Moroccan, 45 Surinamese, 5 Antillean, 3 Mollucan and 63 others (varying from south and north-west Europeans to Africans and Asians). The percentage of migrants on the total school population at the 44 schools varies from 8 to 97%, with an average of 48%. This clearly shows that this is definitely not a national representative sample survey. Not only migrants are over-represented, but also Dutch working class children, who are usually over-represented in migrant classes.

Findings and discussion
The results from our analysis reveal that our intention has only partially been successful. Part of the original 'ethnic' differences in achievement can be seen as an outcome of the social class composition of certain minority groups in the Netherlands. Both teacher perceptions and target levelling towards migrant pupils seem to be social class related in as far as they are originally different for different ethnic groups: minority pupils (possibly most the ones with relatively good capacities) suffer from a social class bias in teacher behaviour. Only in considering this complex interaction, we do not conclude for racist bias in teacher perceptions, which could eventually be done from the original findings.
In last instance our analysis forces us to differentiate between educational effectiveness and school effectiveness. The processes we have tried to indicate, are of relevance in first instance, at the individual pupils' level. When it comes to explaining school differences, they seem to be of no value, eventually just as a suppressor variable.

DIFFERENCES IN RECOMMENDATIONS FOR SECONDARY EDUCATION BETWEEN DUTCH AND IMMIGRANT CHILDREN WITHIN AND BETWEEN SCHOOLS

Lia Mulder, ITS, Nijmegen, The Netherlands

Introduction
From several studies it appears that immigrant children who receive the same recommendations for secondary education compared with Dutch children score lower on achievement tests in grade 8 of primary education. Different explanations for this phenomenon has been given. Most explanations suppose that *within* schools immigrant children receive higher recommendations than Dutch children with the same achievement scores. Reasons mentioned for this are, among others, positive discrimination and compensation for language disadvantage. However, it

1339

might also be possible that the explanation should be sought *between* schools instead of within schools. Immigrant children mostly visit schools with a high proportion of disadvantaged pupils. Through this the mean achievement level of those schools is relatively low. A bright pupil within such a school achieves almost at the same level as an intermediate pupil within a school with a high mean achievement level. Both types of schools will, at the moment of recommendating for secondary education rank order the pupils in such a way that the better pupils will receive a higher recommendation than the less achieving pupils. The consequence of this is that a better achieving pupil within a school with a low mean achievement level receive a higher recommendation compared to an intermediate pupil within a school with a higher mean achievement level, even though the achievement level of both pupils is the same.
The question that will be answered in this paper is whether the explanation for relatively high recommendations for immigrant children should be sought within or between schools.

Method
To answer the question we used the data that were collected in the national evaluation of the Educational Priority Program. From the children that were tested in 1988 for intelligence, arithmetics and Dutch language also the recommendations for secondary education are available. As background variables we used the country of origin and the weight factor as an indicator of social background. To establish the school effects on the recommendations for secondary education, the data were analysed with the multi level program VARCL.

Results
From the analyses it appears that a bit more than 10% of the variance in recommendations can be explained by the school: some schools give higher recommendations than others. For a part these differences between schools can be ascribed to differences in the language and arithmetic scores of the pupils within these schools. After controlling for the individual achievement scores however there still exist differences in recommendations between schools (8% variance explained). After controlling for the mean achievement level of the schools only the mean achievement level for arithmetic contributes to the between school variance. The between school variance now is 7%. The effect of the mean arithmetic level is negative. The lower this mean scores, the higher the recommendations are. This supports the hypothesis that immigrant children relatively receive high recommendations because they visit schools with a low mean achievement level.
However, the possibility that also within schools immigrant children receive higher recommendations can not be excluded. Comparison of the effects of the mean arithmetic level of the school and the effect of ethnic origin of the pupils on the recommendations shows that both effects are almost the same.
So the explanation of relatively high recommendations for immigrant children should be sought between as well as within schools. Schools with a low mean achievement level give relatively higher recommendations compared to other schools as well as schools in general give higher recommendations to immigrant

children than to Dutch children with the same achievement scores and the same social background.

THE NEED FOR COOPERATION OF EDUCATIONAL PRIORITY POLICY AND BASIC ADULT EDUCATION

Gerry J. Reezigt, Institute for Educational Research (RION), Groningen, The Netherlands

Introduction

Official documents on educational priority policy and basic adult education often state that some form of cooperation should exist between both fields of education. One of the practical reasons is an overlapping target group: the parents of disadvantaged children are often the adults considered to be students in basic adult education. This situation argues for at least a minimum of mutual consultation, to prevent an unnecessary overlap of activities.

Another reason for cooperation is that further schooling of parents may contribute substantially to a more successful school career of their children. This holds especially strongly for parents who cannot read or write and for immigrant parents who can't speak the Dutch language. These parents will not be able to read to their children or help them when they start reading and writing in Dutch.

Moreover, a problem sometimes found with illiterate adults is a negative attitude towards school and learning (because of their own unfavourable experiences) that they may pass on to their children.

A third reason for cooperation is that, by exchanging knowledge on various aspects of educating disadvantaged students, a further growth of illiteracy may be prevented. At this moment elementary education still delivers students lacking basic skills, who then have to apply for adult education some years later.

Research problem

In the early stages of the Dutch educational priority policy in the sixties and seventies, cooperation between teachers of elementary schools and adult educators was rather widespread. Moreover, many adult educators actually were elementary school teachers. Teachers, trying to change the unfavourable school career of lower–class children, detected a lack of basic educational knowledge (reading, writing, arithmetic) in their parents. They started courses for these parents, to improve their basic skills, and taught them themselves. In the seventies it became clear that even in the Netherlands as an affluent society there were many parents who could not read and write at an adequate level, even though they finished elementary education and sometimes even some years of secondary education. Elementary school teachers then started literacy courses, and these courses, among others, formed one of the elements of the current curriculum in basic adult education. In the eighties the educational problems of immigrant students and their parents became clear. Immigrant students started to form one of the target groups of the educational priority policy, together with Dutch lower class–students. The problems of their parents were even more severe than the

problems of Dutch illiterate adults, because of their disability to read or write in their own language as well as in the Dutch language.
In the eighties new legislation came into force for educational priority policy as well as for basic adult education. This implied the establishment ofinstitutes for basic adult education and a separation of the formerly spontaneous integration of activities, and also a separate financial background. Many projects involving teaching both parents and children disintegrated. When this situation was discovered, new ways of cooperation were stimulated by policy makers in official documents, but no specific financial support was given.
A research project was started as a part of the national evaluation of the educational priority policy, to find out whether forms of cooperations still existed, or maybe just had started, and also to find out how at low cooperation could be stimulated.

Research activities
To answer the research questions, the following activities were untertaken:
- ducational priority areas where cooperation with basic adult education is taking place;
- describing actual cooperation;
- tracking down ideas and suggestions for the improvement of cooperation.

Results
To find out if any cooperation existed, case studies in 9 educational priority areas were carried out and 22 representatives concerned with either educational priority policy or basic adult education were interviewed.
When cooperation actually takes place it almost always concerns courses for parents to improve their basic skills in language or math, or courses for immigrant parents to learn to speak the Dutch language. Sometimes courses on parent–teacher organizations and ways for parents to participate actively in the school of their children are also offered. Hardly any activities directly concerning prevention of illiteracy were found.
Successful and problematic aspects of cooperation will be discussed in the final paper, as well as suggestions for more substantial forms of cooperation.

EFFECTIVE STRATEGIES FOR ETHNIC MINORITY CHILDREN IN PRIMARY SCHOOLS

Marga G. Weide, COWOG, University of Groningen, Groningen, The Netherlands

Introduction
Ethnic minority children generally perform less well in school than their classmates. Their achievement in primary education is lower and after they finish primary school, they find themselves more often in the lower levels of secondary education.

The lower performance of ethnic minority children has been explained in different ways. Explanations can be summarized in two main visions: the deficit– and the cultural difference–thesis.
The first thesis mainly blames the home environment. According to this thesis ethnic minority children lack at home the stimuli necessary for optimally developing their verbal and intellectual skills. The fact that many of them speak a different language at home than the language used in school, is considered an extra handicap.
The second thesis explains the lower performance of ethnic minority children by the differences between home and school environment, such as a different use of language, different (codes of) behavior, different norms and values, in short a different culture.
Educational strategies aimed at improving (lower class) ethnic minority childrens achievement are, implicitly or explicitly, based on one of these visions.
Compensatory measures (for example, extra language lessons) and strategies aimed at maximizing the time children actually spend on learning, try to compensate for the disadvantaged start of ethnic minority children in school.
Activities such as improving the home–school relation and the intercultural adaptation of educational materials try to bridge the gap between home and school.
For the past twenty–five years in the Netherlands many educational strategies have been tried in programs aimed at raising the acievementlevel of lower class (ethnic minority) children and most of these strategies now can be found in everyday educational practice.
This paper deals with the question which of the different strategies are (most) effective for ethnic minority children.

Method

Sample and variables

To answer the question data collected in more than 600 schools in the context of a large survey (the national evaluation of the Educational Priority Program) are analyzed. The sampling of schools was partly at random, partly stratified, in order to guarantee a sufficient number of lower class (ethnic minority) children in the sample. The design of the evaluation is longitudinal: every two years data are collected in the same schools, in three age–groups. Students are tested by means of standardized tests for intelligence, language and math achievement and data about their social and ethnic background is collected. The principal and teachers of the children fill in a written questionnaire on policy, methods, activities and actual (instructional) behavior. As yet, datacollection took place in 1988 and 1990. For this paper only students who were tested in 1988 in grade 6 (age group 9 to 10) and in 1990 again in grade 8 (age group 11 to 12) are considered. Moreover, the paper only concerns schools with a minimum of 10 and a maximum of 90 percent ethnic minority children and ethnic minority children who were less than three years of age when they arrived in the Netherlands.
At the student level intelligence, language and math achievement, social class and ethnicity are the central variables in the analyses. At the school/teacher level the ethnic composition of the student population is taken into account, apart from the different educational strategies.

Educational strategies can be labeled as compensatory measures, school policy emphasizing academic standards, effective instruction, a good home–school relationship and adapting education to the social and cultural background of pupils.

Analyses

To establish the effectiveness of educational strategies, two 'dimensions' of effectiveness must be taken into account: quality and equity. The quality of a school is measured by the mean achievement level of its pupils. However, if the quality of a school is high, the mean achievement of a specific group of pupils (for example ethnic minority children) still can be relatively low. That is why it is important to look at the equity dimension as well, i.e. the difference in achievement between, in this case, Dutch and ethnic minority children.

The analyses are carried out with the multi–level program VARCL.

Research questions are:

- do schools differ in quality and if so, which educational strategies can explain these differences?
- do schools differ in equity and if so, which educational strategies can explain these differences?

As a first step, the situation in 1988 is analyzed, to establish the quality and equity of the schools and the possible influence of variables at the school–level (f.e. school policy). In the second round of analyses the attainment of pupils (achievement in 1990 controlled for achievement in 1988) is the dependent variable and the analyses focus on the influence of teachervariables.

Results

Preliminary analyses show that schools differ in quality, but hardly in equity.

As for the quality dimension, more specific results are:

- In 1988 the initial variance at school level is 17 percent for language and 12 percent for math achievement; after controlling for (non–verbal) intelligence and social and ethnic back–ground the percentage of unexplained variance at school level is 9 and 7 for language and math respectively;
- In 1990 the initial variance at the school–level is 12 percent for language and 11 percent for math achievement; after controlling for achievement in 1988, intelligence and social and ethnic background, 5 and 7 percent of variance is left to explain at school level for language and math respectively.

This means about half of the initial variance at school level is explained by student variables. Although the variance left may seem small, it still is substantial enough to conclude that it does matter which school a student visits.

The equity–dimension is defined as the difference in achievement between lower–class Dutch pupils and ethnic minority children, controlled for intelligence and former achievement. This difference can be represented by a regression line, a so–called slope. Slopes are modelled for each ethnic group in comparison with Dutch lower class students. When there is variation in slopes, schools differ on the equity–dimension. That means that in some schools there is hardly any difference in achievement level between Dutch lower class students and (a specific group of) ethnic minority children, while in other schools the difference is substantial.

The analyses show that there are only a few slope differences and these differences are small.
Results about the effects of different educational strategies will be presented in the final paper.

VISION, APPROACH AND EFFECTIVENESS IN THE DUTCH EDUCATIONAL PRIORITY PROGRAM

Lud van der Velden, Centre for Applied Research on Education (OCTO), Enschede, The Netherlands

In this presentation some results are reported from a study into the effects of the pedagogical vision of teachers and schools on their educational approach and through this on the cognitive and social–emotional functioning of their at–risk–students. The analysis concentrates on autochtonous at–risk–students from the lower social–economic strata.

The theoretical model
What can schools and teacher do to help at–risk–students? In Dutch education there are three competitive schools of thought ('visions') on which they can base their approach. The first vision says that teachers and schools should use a subject–matter or cognitive oriented approach, in which most efforts are directed at the cognitive development of the children. This is the 'equal opportunity vision'. The second vision says that they should use a child or non–cognitive oriented approach, in which most efforts are directed at the social–emotional well–being of the children. This is the 'reform–pedagogical vision'. The third vision says that they should use a group oriented approach, in which most efforts are directed at the social–economical well–being of the children. This is the 'collective emancipation vision'. In Dutch literature on educational disadvantageness, these competitive visions have played a role in the debate.
It is clear that, theoretically, the idea's of a teacher or school can only have an effect on educational outcomes, if differences in idea's go together with differences in the way pupils are treated. Especially important are those characteristics of the teaching–learning–proces that have an influence on the opportunity to learn and the active learning time of the pupils. An effect of adherence to a particular vision should in one way or another be mediated through variables measuring the day–to–day approach.

The data
The study uses longitudinal–data from 287 schools with at least five autochtonous at–risk–students for which we have testscores on intelligence, language and arithmetic in the first half of grade 6 and testscores for language and arithmetic in the first half of grade 8. Some schools had two or three grade 6 groups, which were treated as if they were stemming from different schools. So, in total we had 305 'schools' (or groups), with 2714 1.25–pupils who were tested twice. For half the groups, the school–appreciation of the pupils in grade 8 was measured with a five factor attitudinal scale. Besides the individual (and aggregated) input and

output scores of the pupils, we have survey–data concerning the schools and the directors and the grade 6 and grade 7 teachers for most groups.

The operationalizations of vision

Several variables measuring the vision were available. An 11–item attitude–scale (Cronbach's alpha=.80) measured the importance teachers attach to cognitive goals, while an 8–item attitude–scale (Cronbach's alpha=.69) measured the importance teachers attach to a cognitive approach. Although the two scales were not strongly correlated (.38), they could be combined into a 19–item scale (Cronbach's alpha=.81) that differentiates between cognitive oriented and noncognitive oriented teachers. Representatives of the group oriented vision could not be identified. When aggregated to the schoollevel, the mean attitude of the teachers differentiates between cognitive oriented and noncognitive oriented schools.

In this way, high scores on each of these vision variables indicates adherence to the 'equal opportunity vision', while low scores indicates adherence to the 'reform–pedagogical vision'.

The operationalizations of approach

A large set of variables measuring the approach was available. Based on (American) empirical results from prior studies, a set of variables was included concerning 'effective school' and'effective instruction' characteristics. Also, some variables measuring other kinds of characterisics were included. Most variables consist of some kind of combination of likert–type items, with reliabilities (Cronbach's alpha) laying between .57 and .92, with most reliabilities laying between .70 and .85. Some variables consisted of only one item, measuring some concrete behavior (or perception of behavior).

The operationalizations of effectiveness

A multilevel analysis, using the VARCL–program, showed that 12% of the variance in language scores and 19% of the variance in arithmetic scores in grade 8 lies between groups. However, half of the differences in outcome between groups were related to differences in input between groups. Controlled for input test scores, only 6% of the language variance and 9% of the arithmetic variance were found to lie between groups (or teachers or schools). This result is congruent with earlier analyses of differences in output scores between schools. Though the posterior means for language and arithmetic are positively correlated (.38), the outcome scores are largely independent.

For the general school–appreciation, and the appreciation of the teachers and the organisation, the between–groups variance amounted some 20% of the total variance. The posterior means for these outcome variables showed high positive mutual correlations. For the appreciation of the class–mates and their own functioning, only 6% of the variance was between groups. These posterior means showed positive correlations (.35) with each other and the other three school–appreciation variables. Controlled for input scores, the between–groups variance stayed the same.

The posterior means for language and arithmetic showed no systematic relations with the posterior means for the school–appreciation variables.

Using the multilevel analysis only as a tool for scoring the groups on each outcome variable, the posterior means of the groups were then exported for use outside the multilevel program. This gave the possibility to inspect these effectiveness or quality scores more thoroughly through the use of descriptive and hypothesis–testing tools available in a general statistical package, such as SPSS–X. In this way, the quality scores could be related to school–, teacher– and classvariables using different uni– and multivariate analyses. This meant also that the multilevel analysis itself could be performed for all 305 groups, and not only for those groups without missing data on each of the school–, teacher– and classvariables.

The relations between vision and approach

From a set of fifteen schoollevel variables, only five were systematically related to one or more vision variables. Adherence to a cognitive vision (for all vision variables) goes together with less cooperation within the team (as reported by the director). Adherence to a cognitive vision (for each vision variable of the director and grade 7 teacher) was negatively correlated with the presence of a remedial teacher and the number of ways the remedial teacher is used. Adherence to a cognitive vision (for the director and the grade 6 teacher) was negatively correlated with the use of a student monitoring system for language and arithmetic.
From a set of 23 variables concerning the grade 6 teacher, 13 were systematically related to one or more vision variables. Negative correlations were found between all vision variables and the importance attached to a close connection between the school and the home–environment and the acceptance of flat language usage. Positive correlations were found for the relations between the vision variables and the use of whole–group–instruction and negative correlations for the use of differentiation (for within–class–ability–grouping and the mastery–learning–model). Adherence to a cognitive vision for the grade 6 teacher was negatively correlated to the percentage of communicative language usage compared to a more traditional scholastic language usage. Negative correlations were also found for the systematic evaluation of language and arithmetic learning resuits by the grade 6 teacher. The use of external tests for arithmetic by the grade 6 teacher, was also negatively correlated to the vision variables of the grade 6 and 7 teacher. The use of minimum goals by the grade 6 teacher, was positively correlated to most vision variables.
From a set of 22 variables concerning the grade 7 teacher, 13 were systematically related to one or more vision variables. These relations resembled the ones found for the grade 6 teacher'sapproach. Adherence to a cognitive vision means less importance attached to the close connection between school and the home–environment, more use of whole–group–instruction, less use of differentiation, more use of minimum goals, less use of a systematic evaluation, less use of remedial teachers and less use of a pupil–record.
Because of the way the vision variables are operationalized, the above mentioned outcomes can also be described as effects of adherence to a non–cognitive vision. In case of a positive relation between the posterior means and adherence to a cognitive vision, adherence to a non–cognitive vision will show a negative relation and vice versa.

1347

The negative correlations between the adherence to a cognitive vision and the systematic use of evaluation or the use of external tests, were a surprise: positive correlations were expected. So it can be stated that although vision and approach are related, these relations are not always the ones we expected.

The relations between vision and effectiveness
None of the vision variables were systematically related to the posterior mean scores for language, arithmetic or one of the school–appreciation variables. Although some interactions existed between the input scores of the pupils and the vision variables, controlling for these interactions did not give a different result.

The relations between approach and effectiveness
For language, the specific use of priority–formation and the presence of a remedial teacher did go together with somewhat lower posterior mean scores of the groups. Schoolbased discussion of pupils with learning problems and the use of a student monitoring system for language were slightly positively correlated to the language scores of the groups. The use of whole–group–instruction by the grade 6 teacher, was lightly positively correlated with the language scores, while the use of differentiation based on within–class–ability–grouping was negatively correlated to the quality. The use of immediate inspection of learning results by the grade 6 and 7 teacher, was slightly negatively correlated to the quality. The percentage of the language–curriculum that was covered during the grade 7 period was sligthly positively correlated with the quality. The use of a pupil–record by the grade 7 teacher, was slightly negatively correlated to the quality.
For arithmetic, the same relations were found as for language. Moreover, the posterior mean scores of the groups for arithmetic were also slightly negatively correlated to the number of ways the remedial teacher was used. The use of minimum goals for arithmetic by the grade 6 and 7 teacher, was slightly positively correlated with the arithmetic scores. For the grade 7 teacher's approach, slightly positive relations were also found for the systematic use of tests, the immediate inspection of learning results, the systematic evaluation of learning results and the percentage of arithmetic–curriculum covered during the grade 7 period.
The school–appreciation scores were not systematically related to the approach variables of the schools or the teachers.

Future analysis
Schools and teachers do make a difference, but there is hardly any evidence that certain kind of visions or approaches are more effective than others. There are some seperate variables that are to some extent related to the language and arithmetic scores, but these only account for some 10% of the differences in outcome on the grouplevel. An analysis will be undertaken, trying to relate the posterior mean scores to several variables at one time. It can be stated that although separate variables do not contribute significantly to quality, the variables in conjunction may be of importance.

AUTHORS INDEX

Name	Chapter	Page
Aalsvoort, G. M. van der	3	375
Aarnoutse, C. A. J.	3	295
Aarnoutse, C. A. J.	5	719
Achtenhagen, F.	7	945, 999, 1035, 1037
Ackerman, T. A.	5	705
Admiraal, W.	6	823
Aguilera, A.	6	908, 911
Aho, L.	2	172
Akker, J. van den	2	99, 153, 189
Akkök, F.	5	731
Allegrezza-Carvoyeur, L. S.	9	1251
Alvarez, P.	5	668
Amelsvoort, J. van	3	396
Anderson, C.	8	1130
Ansellem, M.	3	320
Antoli, B.	6	894
Antunes, H.	9	1243
Ardaç, D.	9	1268
Askar, P.	3	258
Askar, P.	5	731
Assink, E. M. H.	3	285
Assink, E. M. H.	4	644
Atkins, M. J.	4	531, 651
Babanin, L.	4	643
Barendse, P.	8	1199
Barker, P.	4	592
Barrett, E.	6	854
Barrows, H. S.	7	954
Barton, L.	6	854
Bassey, M.	2	213
Bates, I.	9	1241
Becher, T.	8	1063
Behrens, M.	7	1001
Beishuizen, J. J.	3	432, 436
Beishuizen, J. J.	4	633
Benins, A.	6	904
Bennett, N.	3	400
Bennett, N.	6	939, 941
Berberoglu, G.	5	708, 748
Berg, E. van den	2	146
Berg, G. van den	8	1158
Berg, J. van den	7	1020
Berg, R. van den	4	521
Bergen, J. van	5	786
Bergen, Th.	3	396, 396
Bergen, Th.	6	931
Berger, M. P. F.	5	692
Berger, N.	3	387
Bergh, H. van den	2	108
Bergmans, A.	3	295
Berings, I.	9	1286
Bestebreurtje, R.	4	584
Bijtel, J. A. L. van	6	796
Black-Branch, J. L.	9	1276
Blackman, S. J.	4	622
Blom, S.	1	79
Boei, F.	6	931
Boekaerts, M.	3	302, 360, 395
Boekenoogen, M. S.	7	994
Boekhorst, A. K.	4	619
Boekkooi-Timminga, E.	5	698
Bock, R. D.	5	662, 753
Bögels, S. M.	3	245
Bokhove, J.	5	783
Bollen, R.	1	71
Bolzoni, M.	4	569
Bonekamp, L. W. F.	4	538
Boon, J.	7	968
Börger, M. A.	6	924
Bosch, H. van den	8	1201
Bosch, L. van den	5	786
Bos, K. Tj.	1	63
Bos, K. Tj.	6	862
Bosker, R. J.	1	19, 86
Bosker, R. J.	9	1215
Bosman, Y.	7	1004
Bostock, S. J.	4	502
Bossink, C.	6	922
Bouhuijs, P. A. J.	8	1187, 1188
Braaksma, J.	4	577
Brandsma, H. P.	1	8, 69
Brandsma, H. P.	6	844
Brandsma, T. F.	7	1040, 1053
Bratko, D.	5	735
Brekelmans, M.	6	935
Breuer, K.	3	415
Brinke, D. ten	5	774
Brinke, J. S. ten	6	820, 932
Bristow, S.	3	405
Bristow, S.	9	1257
Brouwer, K.	7	964

Brown, A.	7	1001
Brown, A.	8	1168
Bruggencate, G. C. ten	8	1107
Bruijn, E. de	7	1015
Brummelhuis, A. ten	2	206
Bruinsma, C.	6	837
Büchel, F. P.	3	407
Büchel, P.	3	407, 409
Burton, R. M.	1	94
Busato, V. V.	3	435
Cachapuz, A.	6	816
Calderheads, J. C.	6	878
Cardelle-Elawar, M.	3	371
Cardelle-Elawar, M.	6	829
Carleer, G. J.	2	186
Carleer, G. J.	4	434, 631
Carré, C.	6	939, 940
Cate, Th. J. ten	6	796
Cate, Th. J. ten	8	1141
Cavendish, S.	9	1221
Cerveró, F.	6	894
Chartier, D.	3	329
Christiaans, H. H. C. M.	3	401
Claessens, M. P. J. G.	8	1196
Clarke, S.	7	1010
Clement, M.	6	885
Cohen, R.	4	587
Collis, B. A.	4	439, 534, 631
Collins, D.	4	502
Commins, B.	4	525
Confalonieri, E.	9	1245
Connelly, F. M.	2	213
Corporaal, A. H.	6	931
Corte, E. De	3	351, 393, 398, 417
Cox, M.	4	631
Creemers, B. P. M.	1	74, 82
Cremers, L.	4	545
Cremers, L.	6	862
Créton, H.	6	935
Cuijk, H. van	9	1286
Czerniawska, E.	3	345
Daal, M. M.	3	429
Daal, V. H. P. van	3	293
Daly, P.	1	60
Dam, G. T	9	1218
DeAyala, R.	5	686
Dekkers, H.	9	1226
Demirel, M.	3	342
Desforges, C.	3	405
Desforges, C.	9	1257
Diana, I. De	4	595, 598
Diana, I. De	5	728
Diggele, J. van	5	783
Dijk, H. van	3	263
Dijk, W. van	3	354
Dijkstra, S.	3	223, 414
Dirkzwager, A.	4	634
Dochy, F. J. R. C.	3	398, 417
Doets, C.	7	1043
Dolmans, D.	8	1133
Donald, J. G.	8	1145
Dong, Z-P.	4	557
D'Onofrio, A.	8	1172
Doornekamp, G.	4	545, 605
Doughty, P. L.	4	625
Draper, J.	6	813
Driel, C. van	6	820
Driel, J. van	8	1189
Dronkers, J.	9	1271
Dumitrescu, A.	2	131
Dumitrescu, A.	6	856
Dunne, E.	6	939, 940
Dunne, R.	6	871
Eash, M. J.	2	169
Eck, E. v.	9	1218
Eberg, J. W.	2	111
Eberg, J. W.	7	994
Edelenbos, P.	1	69
Eeden, P. van den	2	100
Eggen, T.	5	786
Ehrlich, D.	4	572
Eijkelhof, H. M. C.	3	256
Eldering, L.	9	1282, 1283
Elen, J.	8	1178
Elliott, R. G.	6	834
Elshout, J. J.	3	310, 435, 461
Elshout-Mohr, M.	3	434
Elst, G. van	7	977
Ely, D. P.	4	480
Engelen, R.	5	786
Engelhard, G.	5	741
Erder, A. M.	3	342

Ersoy, Y. 4 483
Ersoy, Y. 6 896
Etelpelto, A. 7 951
Evans, D. 1 55
Fátima Choráo C. Sanches, M. de 3 269
Feijen, C. J. 7 1029
Fergusson-Hessler, M. 3 232
Fernandez Benitez, V. 6 858
Feteris, A. 7 1017
Finegold, M. 4 457
Fischer, D. 6 936
Fitz-Gibbon, C. T. 1 16
Fleming, R. 6 904
Forcheri, P. 3 323
Fox, M. 8 1172
Fraser, H. 6 813
Fraser, B. 6 936
Free, E. L. 8 1114
Frey, K. 2 197
Fricke, R. 4 617
Frijns, P. 8 1193, 1199
Furlong, J. 6 854
Gaag, F. J. van der 4 613
Geban, O. 3 258
Geest, T. M. van der 2 187
Gelderblom, A. 7 980
Geus, W. de 3 298
Geus, W. de 5 744
Gielen, I. 3 351
Giesbers, J. 1 90
Giesbers, J. 6 931
Gijselaers, W. H. 8 1133, 1202
Gillijns, P. 5 786
Glas, K. 5 774
Glendinning, A. 9 1254
Glopper, K. de 3 298
Gmelich Meijling, M. R. 4 490, 613
Goldstein, H. 1 11
Goldstein, H. 5 751
Gómez-Crespo, M. A. 3 381
Gorsky, P. 4 457
Graaff, E. de 8 1187, 1188, 1193, 1208
Grace, G. 6 828
Gravemeijer, K. 2 103
Gray, J. 1 40
Groen-de Jong, C. 3 364
Groot, M. de 3 375
Groot, T. 4 619
Groppo, M. 9 1245
Gruber, H. 7 1038
Grunefeld, H. 8 1069
Guillén, J. 3 390
Gulmans, J. 4 521
Haan, D. M. de 1 75
Haapala. A. 2 172
Hacquebard, A. E. N. 2 165
Halász, G. 9 1275
Hambleton, R. K. 5 681
Hamers, J. 3 407
Hameyer, U. 2 99, 153
Hansen, K-H. 4 487
Harlen, W. 2 175
Harvard, G. 6 871
Heathcote, G. 6 793
Heitmann, G. 8 1126
Hell, J. G. van 4 461
Helmke, A. 3 398
Helsper, E. 4 640
Hendry, L. B. 9 1254
Henneken, O. W. H. 5 744
Hermans, H. 3 422
Herrlitz, W. 2 212
Heuvel-Panhuizen, M. van den 2 103
Heuvel-Panhuizen, M. van den 3 312
Hippe, Z. S. 4 465
Hoeben, W. Th. J. G. 2 210, 212
Hoeven-van Doornum, A. A. 9 1264
Hofman, R. H. 1 46
Holmberg, L. 5 717
Holsbrink-Engels, G. A. 3 243
Hoogstraten, J. 8 1077
Hopkins, D. 1 62
Hough. M. 6 846
Hout, J. F. M. J. van 8 1155
Houtveen, Th. 6 802
Hout-Wolters, B. H. A. M. van

	3	432
Hout-Wolters, B. H. A. M. van	6	791, 930
Hövels, B.	7	1020
Huijsman, Z. C.	8	1081, 1153
Hulst, A. van der	3	357
Hutchison, D.	5	756
Huteau, M.	3	329
Ibrahim, A. B.	1	16
Izard, J.	5	775
Jagt, P. van der	6	837
Jahn, H.	8	1123
Jakowicka, M.	3	348
James, C.	6	878
Jansen, E. P. W. A.	8	1073
Janssen, J.	5	783
Janssen, J.	9	1228
Janssen, P. J.	3	402
Janssen, T.	7	1041
Janssen Reinen, I.	2	202
Järvelä, S.	3	272
Jaspers, J. G. M.	4	560
Jeffcut, P.	8	1081
Jenkins, H.	1	55
Jochems, W. M. G.	8	1084, 1088, 1196, 206
Johnson, S.	1	13
Jong, F. P. C. M. de	3	396
Jong, J. de	5	721
Jong, J. A. de	7	971, 977
Jong, M.J. de	6	865
Jong, R. de	6	806
Jong, W. J. de	8	1088
Joosten, G.	9	1228
Jorge, M.	5	816
Jost, K. L.	4	625
Jung, M. H.	4	617
Jungbluth, P.	9	1264, 1284, 1311
Jurgens, H. J. H.	8	1155
Kaldeway, J.	3	337
Kamp, L. J. Th. van der	5	759
Kamp, M. van der	7	1041
Kanselaar, G.	2	218
Kanselaar, G.	4	560
Karstanje, P. N.	1	42
Karstanje, P. N.	7	1004
Kattenberg, G.	3	285
Kauppi, A.	6	905
Kessels, J.	7	1025
Kesteren, B. J. van	4	534
Keursten, P.	2	143
Kieviet, F.	6	919, 931
Kirschner, P. A.	3	417, 422
Kleintjes, F. G. M.	1	25
Kley, P. van der	9	1286
Kluvers, C.	3	396
Knuver, A. W. M.	1	8
Koetsier, C. P.	6	820
Kohlemainen, P.	8	1111
Kohn, D. I.	3	384
Kok, W. A. M.	3	306
Kok, W. A. M.	4	560
Kommers, P.	4	580
Koning, J. de	5	739
Koning, J. de	7	980
Kontogiannopoulou-Polydorides, G.	2	198
Koolstra, C.M.	5	759
Koopman, P.	8	1100
Kop, J. L.	9	1251
Koper, E. J. R.	4	634
Korthagen, F.	3	337
Korthagen, F.	6	823, 930, 932
Korthals, M.	8	1182
Kovacevic, P.	5	735
Kraemer, J. M.	5	783
Krause, B.	3	267
Kremers, E. J. J.	1	25
Krogt, F. J. van der	7	973
Krykorková, H.	3	249
Kuiper, W. A. J. M.	2	155
Kusá, D.	6	916
Ladhani, A-N.	4	595
Landucci, M.	2	117
Lang, M.	2	200
Langereis, M.	8	1077
Lautrey, J.	3	329
Leal Filho, W. D. S.	2	121
Lebedev, A. N.	5	734
Leemkuil, H. H.	3	325
Leenders, Y.	6	900
Leeuwe, J. F. J. van	5	719
Leiblum, M. D.	4	549
Lera, M.	6	908

Letiche, H. K. 8 1081, 1153
Levy, F. 3 320
Levy, J. 6 936
Lewis, B. 4 468
Lijnse, P. L. 3 256
Lima, A. de 6 880
Linden, E. van der 4 567
Linden, J. van der 4 644
Linden, W. van de 5 657
Lindström, B. 4 632
Lingefjärd, T. 4 647
Liverta Sempio, O. 9 1245
Loarer, E. 3 329
Lockwood, F. 5 660
Lodder, B. J. H. 8 1057
Lodder, B. J. H. 9 1235
Logar-Djuric, S. 5 735
Lombardini, H. E. 2 117
Loo, P. J. E. van de 8 1057
Lopez, A. M. 5 762
Love, J. G. 9 1254
Lowyck, J. 3 415
Lowyck, J. 6 840
Luyten, H. 1 74
Macready, G.B. 5 686
Main, R. G. 3 368
Maltseva, I.V. 5 734
Mandl, H. 7 1038
Manfred, E. 5 685
Mantovani, G. 4 569
Mª Poulussen van Balen, J. 6 858
Marinov, S. 5 711
Markinev, A. V. 5 734
Martens, R. L. 3 429
Martinez, R. 9 1293
Martinkova, V. 2 128
Martiny, U. 8 1060
Mason, G. 6 936
Mason, L. 3 333
Masters, G. N. 5 664
Matthijsen, C. 6 844
McGuigan, L. 2 209
McGuigan, L. 3 261
Meer, L. v. d. 8 1182
Meer, S. van der 8 1133
Meester, M. A. M. 3 422
Meesterberends-Harms, G. J. 7 958
Meijer, P. J. 3 281
Meijer, P. J. 4 563
Meijer, R. R 5 665
Mellink, E. C. 7 1004
Merriënboer, J. van 3 411
Menis, J. 2 213
Meulemann, H. 9 1273
Meyer, J. R. 6 849
Meyer, P. 6 900
Middelbeek, E. 3 422
Miles, S. 6 854
Milikowski, M. 3 310, 435
Miller, L. 4 609
Millett, A. 2 208
Min, F. B. M. 4 490, 613
Minnaert, A. 3 402
Mirande, M. J. A. 4 549
Moelands, F. 5 786
Moerkamp, T. 7 1044
Molenaar, I. W. 5 673
Molfino, M. T. 3 323
Mooij, T. 2 100
Moonen, J. 4 439, 633
Mora, J. G. 6 911
Mora, J. G. 9 1293
Morais, A. M. 9 1243
Morales, M. 5 762, 764
Moreno 5 764
Moskovkina, A. G. 9 1262
Mourik, T. G. C. van 3 245
Mulder, F. 2 165
Mulder, J. W. 8 1117
Mulder, L. 9 1312
Mulder, M. 7 945, 1021
Munn, P. 8 1093
Muraki, E. 5 662
Naylor, A. 6 914
Neber, H. 7 1039
Neut, I. v. d. 4 446
Neuwahl, N. M. E. 3 364
Nicaise, I. 9 1231
Niederer, R. 2 197
Niet, G. 4 613
Nieuwenhuis, A. F. M. 7 991, 1015
Nieuwenhuizen, C 3 302
Nieveen, N. 2 143
Nijhof, W. J. 7 945, 1033

Nguyen-Xuan, A. 3 320
Obel, B. 1 94
O'Brien, J. P. 6 914
O'Connor, J. E. 4 454
Oestermeier, U. 3 236
Ojen, Q. H. J. M. van 1 42
Ojen, Q. H. J. M. van 6 851
Ok, A. 6 809
Olsen, J. 2 184
Olsen, J. 4 609
Onna, B. van 7 960, 1029
Onstenk, J. 7 1046
Oort, F. J. 5 708, 748
Oosterhuis-Geers, J. 8 1065
Otten, R. 3 395
Oquist, O. 2 141
Oud, J. H. L. 5 719
Pan, H. 1 11
Pan, H. 5 751
Pavia, F. 6 894
Paykoç, F. 6 809
Peisa, S. 6 905
Pelgrims, G. 3 407, 410
Pelgrum, W. J. 1 75
Pelgrum, W. J. 2 195, 202
Perez-Eccheverria, M. P. 3 378
Petrikás, A. 1 31
Pieters, J. M. 6 791
Pieters, M. 2 115
Pijnenborgh, F. 3 250
Pilot, A. 4 577
Plomp, Tj. 2 193, 195, 202, 213
Plomp, Tj. 4 534
Pluijm, J. van der 9 1259
Podolskij, A. E. 8 1181
Poelmans, P. 3 429
Pool, I. 7 977
Poorthuis, G. M.T. 3 306
Popham, W. J. 5 713
Poschner, M. T. 8 1165
Pouw, C. L. M. 8 1114, 1182
Powney, J. 2 213
Powney, J. 8 1161
Pozo, J. I. 3 378, 381
Prucha, J. 2 125
Puchová, I. 3 276
Qualter, A. 2 209
Raban, B. 2 208
Raffe, D. 7 985
Ramaekers, G. W. M. 8 1057
Rasbash, J. 1 11
Rasbash, J. 5 751
Reezigt, G. J. 1 34
Reezigt, G. J. 9 1314
Reimerink, H. B. 4 613
Reitsma, P. 3 289, 298
Reitsma, P. 5 744
Reitsma, P. 6 874
Reisby, K. 2 135
Renkl, A. 7 1038
Renyier, M. 2 117
Resing, W. C. M. 3 387
Reynolds, D. 1 62
Reynolds, L. 4 572
Rezendes, G. 5 685
Riet, K. van 't 4 493
Rijns, Y. 8 1088
Riley, K. A. 1 58
Ris, Th. P. J. 4 615
Robijns, M. C. 9 1279
Roebuck, M. 4 529
Roeleveld, J. 9 1279
Rodriguez, D. 3 390
Rodriguez, R. 6 936
Roes, M. G. 2 148
Rojero, F. F. 2 117
Roode, F. de 5 774
Ropo, E. 2 133
Ropo, E. 6 826
Ruijs, G. P. C. M. 7 1050
Russell, T. 2 206, 209
Russell, T. 3 261
Sagdullaev, A. A. 9 1262
Sainsbury, M. 5 779
Saldaña, D. 3 390
Salonen, U. 2 172
Samarapungavan, A. 3 435
Sanders, M. 6 900
Sanz, A. 3 378
Saroyan, A. 8 1145
Schadé, E. 8 1141
Schagen, I. 5 756, 776
Schaick, P. van 4 513

Schaick Zillesen, P. G. van 4 490, 493, 613
Scheerens, J. 1 3, 80, 85, 86
Schmidt, H. 8 1133
Schmidt-Hackenberg, B. 7 989
Schmitt, M. 3 285
Schoenmaker, J. 4 468
Schott, F. 3 416
Schott, F. 8 1181
Schouten, A. 4 637
Schranhorst, U. 3 407, 411
Schreurs, M. L. J. J. 7 947
Schuurs, U. 2 108
Schuurs, U. 5 721
Sciarone, A. G. 3 281
Sciarone, A. G. 4 563
Scaratti, G. 9 1245
Seegers, G. 3 302, 360
Segers, S.R. 5 772
Sijtsma, K. 5 665
Sijtstra, J. 5 784
Simons, P. R. J. 3 395, 396, 411, 417
Simons, R-J. 3 223
Siregar, N. 6 867
Sheridan, B. E. 5 671
Shucksmith, J. 9 1254
Slettenhaar, H. K. 8 1117
Sliepen, S. E. 6 874
Sligte, H. W. 4 509
Smeets, D. 2 165
Smid, H. J. 8 1088
Smith, K. 5 772
Snippe, J. 2 151
Sola, I. 3 390
Spencer, E. P. 1 5
Spoon, P. 4 446
Squarcy, K. 5 735
Staa, A. van 7 968
Staessens, K. 6 891
Stam, H. van der 8 1120
Stanchev, I. 4 439
Stanley, S. J. 2 138
Steiner, G. 3 239
Stenström, M. L. 9 1248
Stoel, W. G. R 6 844
Stokking, K. M. 1 49
Stolk, H. 3 425
Stoutjesdijk, E. 3 436
Streumer, J. N. 7 945, 1009, 1017
Strum, J. 2 212
Struyf, E. 3 351
Sultana, R. G. 9 1238
Sumner, R. 8 1102
Taconis, R. 3 232
Tartwijk, J. van 6 935
Taylor, W. 6 813
Terekhina, A. 4 641
Tergan, S-O. 3 236
Terlouw, C. 8 1114, 1175
Terwel, J. 2 99, 100
Thompson, J. 4 525
Tiemessen, W. J. M. 8 1117
Tillema, H. H. 6 919
Timmer, E. 8 1091
Timmer, J. 4 544
Tobi, H. 5 695
Toluk, Z. 6 896
Tomic, W. 2 168
Trumper, R. 2 225
Tuijnman, A. 2 206
Tymms, P. B. 1 37
Ulbricht, K. R. G. 5 767
Underwood, G. 9 1221
Underwood, J. 9 1221
Valcke, M. M. A. 3 429
Valk, A. E. van der 3 256
Vandenberghe, R. 6 882, 885, 888, 891
Vedder, P. 9 1282, 1283
Veen, A. M. 9 1279
Veen, C. H. J. van 8 1081, 1153
Veen, J. T. van de 8 1107
Veen, R. van de 8 1091
Veen, W. van 4 446
Veenman, M. V. J. 3 435, 461
Veenman, S. 6 900
Veenstra-Strijland, T. 2 165
Vegt, A. L. van der 6 802
Velden, L. v. d. 9 1318
Velden, R. van de 9 1235
Velden, R. K. W. van der 8 1057
Vendrig, A. S. A. 7 971

Verhagen, P. W. 4 518, 584
Verloop, N. 6 926, 934
Vermeulen, C. 1 19
Vermeulen, C. 6 862
Vermeulen, M. 1 42
Vermeulen, M. 8 1096
Vermunt, J. D. H. M. 7 983
Vermunt, J. D. H. M. 8 1079
Verschaffel, E. 3 351
Versloot, A. M. 7 971
Verwey, A. 8 1088
Vila, L. 9 1293
Vinjé, M. P. 5 785
Vinke, D. 8 1084
Visscher, A. J. 1 63
Visser, H. J. 1 42
Vlas, H. 4 545
Voeten, M. J. M. 3 295
Voeten, M. J. M. 9 1264
Vogelzang, F. 4 446
Volman, M. 9 1218
Volpe, R. 6 904
Volman, C. 5 728
Voncken, E. 7 1004
Voogt, J. M. 2 184, 188
Vooijs, M. W. 5 759
Voronel, Y. 9 1289
Vos, H. 3 228
Vos, H. J. 5 725
Vos, M. 1 28
Vouten, M. J. M. 5 719
Vries, S. de 4 580
Vrugt, A. 8 1077
Vu, N. V. 7 954
Walker, D. F. 2 163
Walraven, M. 3 289
Walsh, A. 2 208
Wang, J. W. 4 541
Warmerdam, J. 7 973
Watson, D. M. 4 441
Weerden, J. van 5 783
Weide, M. G. 1 34
Weide, M. G. 9 1315
Werf, G. van der 9 1296
Wesselingh, A. 9 1215, 1286
Westbury, I. 2 182
Weston, P. 2 179
Wetterling, J. 4 553
Wieringen, F. van 1 88
Wiggers, J. A. 7 991
Wierstra, R. F. A. 2 159
Wijngaarden, K. van 8 1138
Wilbrink, B. 5 701
Wilbrink, B. 8 1149
Wignall, R. 1 53, 67
Willems, J. 8 1138
Wise, S. L. 5 689
White, J. 6 800
Whitty, G. 6 854
Witziers, B. 1 21
Woodhouse, G. 1 11
Woodhouse, G. 5 751
Woord, J. van der 8 1208
Woude, J. van der 4 506
Wubbels, Th. 6 820, 823, 930, 936
Yang, M. 1 11
Yang, M. 5 751
Yildiz, R. 4 651
Yontar, A. 3 253
Yrjönsuuri, R. 3 317
Zarevski, P. 5 735
Zeeuw, G. de 4 509
Zhang, J-P. 4 557, 602
Zhu, Z. 4 598
Zimowski, M. F. 5 753
Zinn, B. R. 3 384
Zinn, B. R. 9 1289
Zolingen, S. J. van 7 960
Zuidema, J. 4 640
Zuilichem, J. van 7 983
Zysno, P. V. 5 680
Zwart, W-J. 4 498
Zwietering, M. H. 4 493

BEST COPY AVAILABLE

Dutch Educational Research Association

1358

University of Twente
university for technical and social sciences